THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Correlated Materials
Models & Methods

Hugo U. R. Strand

Department of Physics
University of Gothenburg
Gothenburg, Sweden 2013

Correlated Materials – Models & Methods
Hugo U. R. Strand
ISBN 978-91-628-8643-1
http://hdl.handle.net/2077/32118

Department of Physics
University of Gothenburg
SE-412 96 Gothenburg, Sweden
Telephone: +46 (0)31–786 0000

Cover: Bethe graph with coordination number $z = 3$
and the Gutzwiller wave-function Ψ_G.

Written in Emacs, and typeset in LaTeX;
figures created using Python, Matplotlib and Ti*k*Z.

Printed by:
Ale Tryckteam
Bohus, Sweden 2013

Correlated Materials – Models & Methods

Hugo U. R. Strand
Department of Physics
University of Gothenburg

Abstract

This thesis encompass a series of studies on methods and models for electron systems with local interactions, relevant for correlated materials. The first study focus on the canonical model for local correlation, the Hubbard model. Using dynamical mean field theory, the critical properties of the finite temperature end point of the metal insulator transition are determined. The issue of computing real frequency spectral functions is also addressed through the development of the novel method, distributional exact diagonalization.

Next topic is the multiband Gutzwiller variational method, for which an efficient solver is presented, applicable to realistic d-electron models when accounting for lattice symmetries. The solver is applied to the iron based superconductors FeSe and FeTe, where the Hund's coupling is found to drive orbital differentiation in the correlated parent state.

A central issue is how to model the local Coulomb interaction. Imposing rotational invariance on the complete set of d-states results in the Slater-Condon interaction, to be compared with the simpler Kanamori interaction, that is shown to be a Laporte-Platt degenerate point of the former. The derivation of a minimalistic form for the Kanamori interaction in terms of density-density, total spin, and total quasi-spin operators enables an exact parametrization of the Slater-Condon interaction in terms of the Kanamori parameters.

The additional interactions contained in the Slater-Condon form are identified as higher order multipole scattering, and the parametrization enables a direct study of the effect of these interaction processes. The multipole scattering is found to drive charge disproportionation and valence-skipping for a subset of multipole active d-band fillings, and raises the question whether such multipole effects are manifested in real materials.

Keywords: correlated electrons, Hubbard model, multiband Hubbard model, dynamical mean field theory, exact diagonalization, Landau theory, Mott metal-insulator transition, Gutzwiller method, many-body point group symmetry, iron chalcogenides, Kanamori interaction, Slater-Condon interaction.

List of Publications

This thesis consists of an introductory text and the following papers:

I **The Dynamical Mean Field Theory phase space extension and critical properties of the finite temperature Mott transition**
Hugo U. R. Strand, Andro Sabashvili, Mats Granath, Bo Hellsing, and Stellan Östlund
Phys. Rev. B **83**, 205136 (2011)

II **Distributional exact diagonalization formalism for quantum impurity models**
Mats Granath and Hugo U. R. Strand
Phys. Rev. B **86**, 115111 (2012)

III **Efficient implementation of the Gutzwiller variational method**
Nicola Lanatà, Hugo U. R. Strand, Xi Dai, and Bo Hellsing
Phys. Rev. B **85**, 035133 (2012)

IV **Orbital Selectivity in Hund's metals: The Iron Chalcogenides**
Nicola Lanatà, Hugo U. R. Strand, Gianluca Giovannetti, Bo Hellsing, Luca de' Medici, and Massimo Capone
Phys. Rev. B **87**, 045122 (2013)

V **Local correlation in the d-band, Slater-Condon vs. Kanamori**
Hugo U. R. Strand, Nicola Lanatà, Mats Granath, and Bo Hellsing
(in manuscript)

VI **Valence skipping and effective negative-U in the d-band from repulsive local Coulomb interaction**
Hugo U. R. Strand
(in manuscript)

Papers not included in this thesis:

I **Time-dependent and steady-state Gutzwiller approach for nonequilibrium transport in nanostructures**
Nicola Lanatà, and Hugo U. R. Strand
Phys. Rev. B **86**, 115310 (2012)

II **Discretized Thermal Green's Functions**
Mats Granath, Andro Sabashvili, Hugo U. R. Strand, and Stellan Östlund
Ann. Phys. **524**, No. 3-4, 147-152 (2012)

Specification of my contributions to the appended papers:

I I wrote the entire paper except the section on the periodized Matsubara Green's functions and Iterated Perturbation Theory (IPT). I did the implementation and testing of the Exact Diagonalization code and the calculations, except the one based on IPT.

II I took part in developing the method, made a reference implementation, performed the CT-QMC calculations, and assisted in the writing.

III I implemented the algorithm with support from Nicola Lanatà, and did the calculations.

IV I implemented the code for the Gutzwiller calculation including the construction of the tetragonally symmetrized Gutzwiller wave-function basis. I also did the model calculations.

V I planned the project, did the derivations, calculations, and wrote the paper.

VI I am the sole author.

To my late mother

Contents

Chapter 1

Introduction

This thesis is composed by an introduction and a number of publications. The purpose of the introduction is to give a more in depth description of the applied models and methods, and a short summary of the main findings. Every chapter address a specific topic, and the reader interested in a particular subject is encouraged to selectively read the the most relevant chapters. As indicated by the thesis title *Correlated Materials – Models & Methods* the basic motivation for this work is the family of correlated materials, in particular materials where local electron-electron correlations play a central role. Theoretically the field is largely based on building effective models that capture the essential physics of these materials, and the subsequent solution of such models using many-body methods. The presentation is selective, treating only the classes of models applied in the papers, namely the single band Hubbard model and its d-electron multi-band generalization. On the method side, single band Dynamical Mean Field Theory (DMFT) and multi-band Gutzwiller variational methods are treated. It is the hope of the author that future students entering the field will benefit from these presentations. Intentionally the treatments are condensed, and aims at being self-consistent.

In terms of achievements in research the main contributions of the appended papers to the field are in my humble opinion: *Paper I* – The determination of the DMFT critical exponents of the finite temperature end point of the metal insulator transition in the paramagnetic single band Hubbard model. *Paper II* – The development of the novel distributional exact diagonalization impurity solver working directly on the real axis. *Paper III* – The development of an efficient solution method for the multiorbital extension of

the Gutzwiller method able to treat general local interactions. *Paper IV* – The Hund's exchange is shown to play an important role for the orbital decoupling in the anomalous high-T_c iron chalcogenide superconductors FeSe & FeTe. *Paper V* – The comparison of the two interaction models for five band d-electron systems, the approximate Kanamori interaction, and the full rotationally invariant Coulomb interaction. *Paper VI* – The analysis of the higher order multipole scattering in the d-electron Coulomb interaction and its role in driving charge disproportionation. My contribution to these developments are specified separately for each paper on page vi.

My work has to a large extent been focused on implementing, testing and performing numerical calculations. During the first years I implemented the standard exact diagonalization algorithm, with both full diagonalization and the Lanczos method, and later the distributional-exact diagonalization method. Another big project has been to implement the new Gutzwiller multi-band solver, and the point group symmetry reduction of many-body operators, enabling calculations of d-band models with crystal fields. Apart from coding much effort has gone in to the reduced matrix element analysis of rotationally invariant d-band interactions.

The introductory text is structured in the following way; In Chapter 1 the non-relativistic atomic and electronic problem is introduced in terms of the Schrödinger equation, and the ab initio treatment of the problem in terms of Density Functional Theory (DFT) is briefly discussed. The potential use of DFT results as a starting point for constructing low energy effective models is outlined, and the chapter is closed by giving some tentative experimental examples of correlation effects, relevant to the appended papers. From the general interacting electron model we arrive at the canonical single band Hubbard model in Chapter 2 and some of its general features are touched upon. Generalizing to several bands in Chapter 3 yields a more involved Coulomb interaction, and its rotationally invariant form is dissected maliciously. Then we turn to many-body methods for solving interacting models in Chapter 4 where Dynamical Mean Field Theory is introduced, accompanied by a number of impurity solvers in Chapter 5. Another many-body method, the Gutzwiller variational approach is presented in Chapter 6. The multi-band formalism and our new solution method is outlined. To make multi-band calculations feasible one need to take wave-function symmetries in to account. A computational procedure for imposing lattice point group symmetries is presented in Chapter 7. Finally Chapter 8 summarizes the appended papers and point out the main findings. The introductory text is in part an adaptation of the author's Licentiate thesis [1], in particular Chapters 1, 2, 4, and 5.

1.1 Ab initio electronic structure

The electronic structure on the atomic scale is central for the understanding of the macroscopic properties of solids. Determining the electronic structure is a genuine many-body problem involving about 10^{23} electrons per $1\,\mathrm{cm}^3$ in bulk materials [2]. As inherent quantum objects and spin one half fermions the electrons must be described using quantum mechanics and Fermi-Dirac statistics. Without accounting for relativistic effects, the wave-function Ψ of the system is the solution of the Schrödinger equation

$$i\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle \tag{1.1}$$

with the Hamiltonian [3]

$$\begin{aligned}\hat{H} = &\sum_\alpha \frac{\mathbf{P}_\alpha^2}{2M_\alpha} + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \\ &+ \sum_i \frac{\mathbf{p}_i^2}{2m_e} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \\ &= \hat{T}_{\text{nucl}} + \hat{H}_{\text{nucl-nucl}} + \hat{T}_{\text{el}} + \hat{H}_{\text{el-el}} + \hat{H}_{\text{nucl-el}}\,, \end{aligned} \tag{1.2}$$

where $\mathbf{r}_i$, $\mathbf{p}_i$ and $\mathbf{R}_\alpha$, $\mathbf{P}_\alpha$ are the positions and momenta of the electrons and atomic nuclei respectively, m_e is the electron mass, and M_α are the masses and Z_α the charges of the nuclei. Due to the large separation in mass between the atomic nuclei and electrons $M_\alpha \gg m_e$, the nuclei can to a good approximation be regarded as stationary when solving the electronic problem. This is the so called Born-Oppenheimer approximation [4], which reduces the electron-nuclei interaction to a constant background potential $V(\mathbf{r})$ acting on the electrons, replacing the terms $\hat{T}_{\text{nucl}}$, $\hat{H}_{\text{nucl-nucl}}$ and $\hat{H}_{\text{nucl-el}}$ in Eq. (1.2).

The second part of the 20th century saw great theoretical and numerical advances in methods for solving the interacting electron problem of Eq. (1.2). In 1964 Hohenberg and Kohn published their seminal paper, proving that the ground state energy of the interacting electron system is a functional of only the electron density $\rho(\mathbf{r})$ [5]. This celebrated Hohenberg-Kohn theorem is the basis of Density Functional Theory (DFT) [4], an extremely versatile method with unprecedented predictive power, for which Valter Kohn received the Nobel Prize in Chemistry in 1998. However the existence of a density dependent ground state energy functional does not solve the interacting electron problem, but it paves the way for developing extremely good approximations.

The central step making DFT a powerful computational tool is the mapping of the electron many-body system to a non-interacting Kohn-Sham reference system $\hat{H}_{\mathrm{KS}}$ with a density dependent single-particle effective potential [6]. The total energy can then be written as the sum of the kinetic energy of the non-interacting reference system, the classical electron-electron Coulomb interaction energy, and a correction called the exchange-correlation energy E_{xc}. The correction term E_{xc} includes the error in the kinetic energy approximation using the Kohn-Sham reference system, and all quantum mechanical effects of electronic exchange and correlation. The corresponding exchange-correlation potential $V_{\mathrm{xc}}(\mathbf{r}) = \delta E_{\mathrm{xc}}[\rho]/\delta\rho(\mathbf{r})$, then enters naturally in the effective potential of the Kohn-Sham system and enables DFT calculations in terms of a non-interacting electron system in a self consistent effective potential.

The catch is that the exact form of $V_{\mathrm{xc}}(\mathbf{r})$ is unknown and one must resort to approximations. Two standard approaches are the local density approximation (LDA) [6], where E_{xc} is approximated in every point by the exchange correlation energy of the homogeneous electron gas with a density equal to the local density, and the generalized gradient approximation (GGA) [7], where also the gradient of the electron density is taken into account. Amazingly, these simple approximations of E_{xc} gives DFT predictive power of material properties for a vast set of compounds both metals and insulators.

Instead of settling down content with the state of things, the basic instinct of any scientist is of course to start exploring compounds that can not be described by DFT methods. It turns out that there are compounds such as CoO and $LaTiO_3$ where DFT fails capitally, predicting metallic ground states while in reality the compounds are good insulators [3]. These insulators are denoted as Mott-insulators and belong to a larger class of so called correlated materials containing partially filled bands with transition metal d-character. The aim of this thesis is to investigate theoretical models and methods in order to understand and predict properties of this material class. For the non-chemist an overview of the periodic table and the transition metals are shown in Fig. 1.1.

Transition metal d-states are generally rather localized, which enhances the local electron-electron interaction effects and eventually invalidates the standard exchange-correlation approximations LDA and GGA [8]. Apart from the Mott-insulators there are also metallic compounds on the verge of a Mott transition, where the local correlation gives effective electron mass renormalizations [9], not captured by DFT. Still the correlation effects are limited to a small energy range of a couple of electron volts around the

Fermi level. As DFT still gives a proper description on higher energy scales the DFT result can be used as a starting point for constructing low energy effective models for correlated systems.

To build a low energy effective model of the interacting electron system requires in principle explicit knowledge of the system's many-body wave-function. Unfortunately only the single-particle states of the Kohn-Sham representation is available from a DFT calculation. We will therefore succumb to what every DFT purist would consider as a cardinal sin, and regard the Kohn-Sham reference system as a "real" electron system. Or, put in more kind words, we replace the non-interacting Kohn-Sham reference system with an interacting one.

The solution of the Kohn-Sham system from DFT provides single-particle Bloch wave-functions $\psi_{\mathbf{k}n\sigma}(\mathbf{r})$ and their dispersion $\epsilon_{\mathbf{k}n\sigma}$, where $\mathbf{k}$ is the wave-vector, n the band index and σ the spin z-component. In second quantized form the non-interacting Kohn-Sham Hamiltonian can be expressed as

$$\hat{H}_{\mathrm{KS}} = \sum_{\mathbf{k}n\sigma} \epsilon_{\mathbf{k}n\sigma} c^{\dagger}_{\mathbf{k}n\sigma} c_{\mathbf{k}n\sigma} \,, \tag{1.3}$$

where $c^{\dagger}_{\mathbf{k}n\sigma}$ creates an electron in the state $\psi_{\mathbf{k}n\sigma}(\mathbf{r})$. The Hamiltonian $\hat{H}_{\mathrm{KS}}$ can be used as a starting point for developing simplified interacting electron models for the system in question. The approach presented here is inspired by the derivation of Hubbard [11].[1]

Let us assume that the narrow d-bands going into the model are well separated from any other bands in the system. In this case the band index n can simply be restricted to the d-subset. The number of required bands equals the number of atomic orbitals of the valence sub-shell, resulting in five spin full bands for transition metal d-states. It is possible to transform the Bloch states $\psi_{\mathbf{k}n\sigma}$ to a set of real-space Wannier orbitals $\psi_{\mathbf{R}n\sigma}(\mathbf{r})$ [12]

$$\psi_{\mathbf{R}n\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \psi_{\mathbf{k}n\sigma}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{R}} \,, \tag{1.4}$$

each centered at the atomic nuclei positions $\mathbf{R}$. Using the electron creation and annihilation operators $c^{\dagger}_{\mathbf{R}n\sigma}$ and $c_{\mathbf{R}n\sigma}$ of the Wannier orbitals, $\hat{H}_{\mathrm{KS}}$ can be written in the tight-binding real-space representation

$$\hat{H}_{\mathrm{KS}} = \sum_{\mathbf{R}\mathbf{R}'} \sum_{n\sigma} t^{n\sigma}_{\mathbf{R}\mathbf{R}'} c^{\dagger}_{\mathbf{R}n\sigma} c_{\mathbf{R}'n\sigma} \,, \tag{1.5}$$

[1] Although due to the lack of any Kohn-Sham system at the time he had to resort to using the result of a Hartree-Fock calculation.

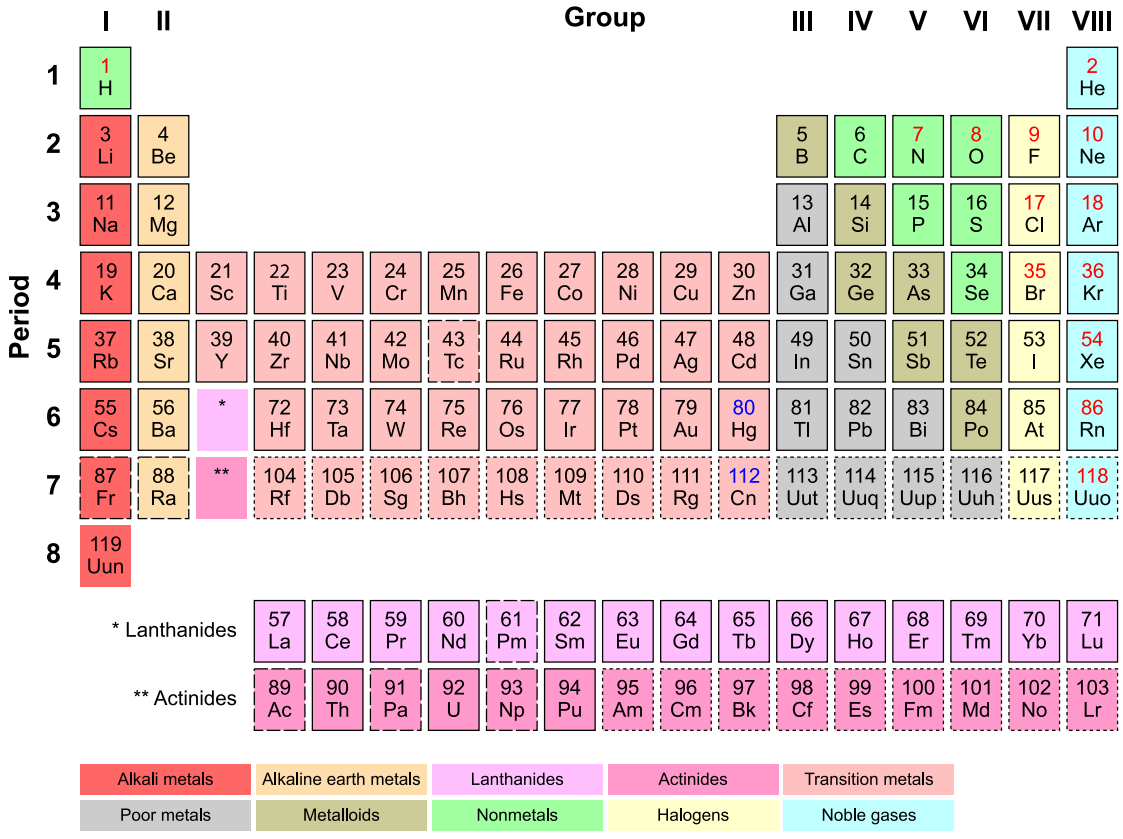


Figure 1.1: The periodic system of elements, adapted from [10].

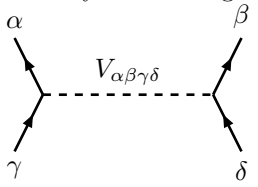
where the single-particle hopping matrix elements are given by [13]

$$t^{n\sigma}_{\mathbf{R}\mathbf{R}'} = \frac{1}{N}\sum_{\mathbf{k}} \epsilon_{\mathbf{k}n\sigma} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \,. \tag{1.6}$$

Further, the two body electron-electron interaction $\hat{H}_{\text{el-el}}$ can be reintroduced in second quantization form as [14]

$$\hat{H}_{\text{el-el}} = \sum_{\alpha\beta\gamma\delta} (\alpha\beta|r_{12}^{-1}|\gamma\delta) c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma = \sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta}\, c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma \,, \tag{1.7}$$

where α, β, γ and δ are super indices containing $\mathbf{R}$, n and σ. The quartic term in Eq. (1.7) corresponds to the two body vertex diagram

$V_{\alpha\beta\gamma\delta}\, c^\dagger_\alpha c^\dagger_\beta c_\delta c_\gamma =$ 

, (1.8)

where the interaction matrix element $V_{\alpha\beta\gamma\delta}$ is given by the Coulomb integral between the respective Wannier wave-functions

$$\begin{aligned} V_{\alpha\beta\gamma\delta} &= (\alpha\beta|r_{12}^{-1}|\gamma\delta) = (\mathbf{R}_i n\sigma, \mathbf{R}_j m\sigma'|r_{12}^{-1}|\mathbf{R}_k n'\sigma, \mathbf{R}_l m'\sigma') \\ &= \iint d\mathbf{r}_1 d\mathbf{r}_2\, \bar{\psi}_{\mathbf{R}_i n\sigma}(\mathbf{r}_1)\bar{\psi}_{\mathbf{R}_j m\sigma'}(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_{\mathbf{R}_k n'\sigma}(\mathbf{r}_1)\psi_{\mathbf{R}_l m'\sigma'}(\mathbf{r}_2) \,. \end{aligned} \tag{1.9}$$

Thus our final effective model Hamiltonian $\hat{H}_{\text{eff}}$ is defined using $\hat{H}_{\text{KS}}$, $\hat{H}_{\text{el-el}}$ and a double-counting term $\hat{H}_{\text{DC}}$ [15] that is used to remove the local part of the electron-electron interaction already included in $\hat{H}_{KS}$ by DFT in a mean-field fashion

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{KS}} + \hat{H}_{\text{el-el}} - \hat{H}_{\text{DC}} \,. \tag{1.10}$$

With this Hamiltonian as a starting point, it is straightforward to work out low energy effective models for materials with strong correlation. Among all possible models the minimal model is the canonical Hubbard model that is the subject of Chapter 2. The generalization to multi-band models for transition metal d-electrons is the subject of Chapter 3.

In order to capture the correlation physics such as Mott-insulating behavior and other exotic phenomena generated by the strong local correlation,

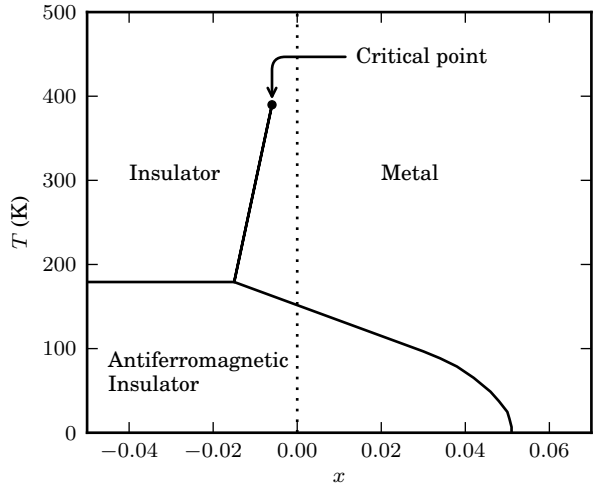


Figure 1.2: Experimental phase diagram of V_2O_3 as a function of Cr and Ti doping x, $(V_{1-x}M_x)_2O_3$ with M = Cr for $x < 0$ and M = Ti for $x > 0$, results from Ref. [8].

these models can in turn be solved with more advanced many-body methods. In this thesis we are going to dwell on two such methods, the Dynamical Mean Field Theory discussed in Chapter 4 and the Gutzwiller variational method treated in Chapter 6.

1.2 Correlation effects

As mentioned, strong local correlations can drive a system into a Mott-insulating state. However this depends on the relative strength of the local Coulomb interaction usually denoted by Hubbard U and the kinetic energy of the electrons scaling with the band-width W of the d-band. The ratio U/W can be controlled by crystal strain induced by, e.g. external pressure or iso-electronic doping, as in the classical example of Cr and Ti doped V_2O_3, where Ti doping acts as an effective positive pressure and Cr doping as negative pressure. At room temperature V_2O_3 is a paramagnetic metal, but a small Cr doping is sufficient to turn it into a Mott-insulator [8], see Figure 1.2. At higher temperature, $\sim$ 400 K, the first order metal to insulator transition ends in a critical point.

As in any first order transition, the metal-insulator transition displays a hysteresis. In a beautiful experiment Limette *et al.* [16] performed conductivity measurements on Cr doped V_2O_3 in a pressure cell and was able to both map out the hysteresis region and determine the critical exponents of the critical end-point. Interestingly the phase diagram as a function of temperature and pressure, or equivalently U/W, has the same qualitative features as the phase diagram of the paramagnetic single band Hubbard model [17]. In Paper I we investigate the corresponding critical end point, confirming that the critical exponents of the Hubbard model agree with those observed in V_2O_3.

It is not only the immediate vicinity of a Mott-insulating state that makes local correlation effects important. There are also many metallic systems where local correlation plays an intricate rôle. Prominent examples are the well known anomalous superconducting iron based compounds. In order of degree of correlation these are the iron pnictide families of 1111 materials (e.g. LaFeAsO), 122 materials (e.g. $BaFeAs_2$), 111 materials (e.g. LiFeAs), and the most strongly correlated chalcogenides FeSe and FeTe [18].

In these systems not only the strength of the direct Coulomb interaction U, but also the Hund's rule exchange coupling J is important [9]. The local exchange coupling J stabilize the metallic state but also drives orbital differentiation and the formation of local moments intermixed with itinerant electrons. Due to their peculiar behavior, this class of compounds are referred to as Hund's metals [19]. In Paper IV we employed the multi-band Gutzwiller formalism, of Chapter 6 and the point group symmetrization of Chapter 7, to study the detailed role of J in the orbital differentiation of the more strongly correlated chalcogenides FeSe and FeTe.

Apart from Mott-insulators and Hund's metals, local correlation also drives a multitude of different orderings, through spontaneous symmetry breaking. Well known examples are of course the plethora of different magnetic orderings, ferromagnetic, antiferromagnetic, incommensurate antiferromagnetic, and different type of glassy magnetic phases. Spin however is only one of the available degrees of freedom in which transition metal atoms can undergo symmetry breaking. Other possibilities are, orbital and charge order, and there is a particularly alluring form of charge order where the transition metal valence disproportionates as, $2d^n \rightarrow d^{n-1} + d^{n+1}$. Well known examples of this type of charge disproportionation are the iron based complex oxides containing nominal valence states Fe^{4+} with four electrons in the d-band.

Mössbauer spectroscopy has been used to analyze the distribution of valence states, making charge disproportionation transitions directly evident.

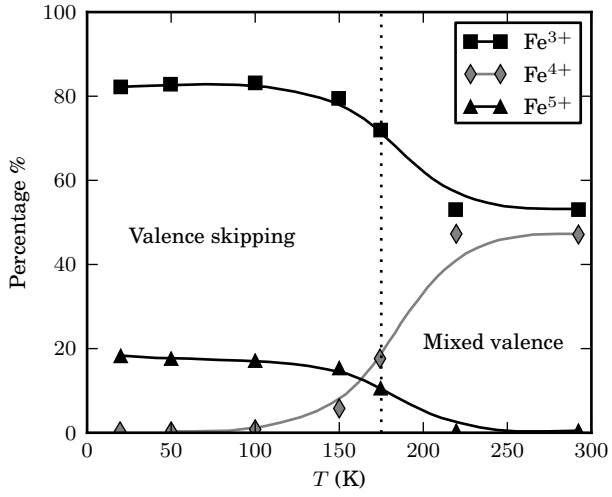


Figure 1.3: Fe valence distribution as a function of temperature in $La_{1/2}Ca_{1/2}FeO_{3-\delta}$ displaying a charge disproportionation transition at $T_{CD} \approx 175\,\mathrm{K}$, data from Ref. [20].

An example of the transition in $La_{1/2}Ca_{1/2}FeO_{3-\delta}$ is shown in Fig. 1.3. With a nominal $Fe^{3.5+}$ valence the system has a mixed-valence state of Fe^{3+} and Fe^{4+} at high temperature, but below the charge disproportionation transition temperature $T_{CD} \approx 175\,K$ the system turns in to a valence-skipping state containing only Fe^{3+} and Fe^{5+} valence states [20].

Usually the charge disproportionation transition is first order and accompanied by magnetic ordering, and often a structural lattice deformation. So it is tempting to assert that charge disproportionation is driven by magnetism and/or lattice distortions. Interestingly there is one system that disproves this assertion. The layered perovskite $Sr_3Fe_2O_7$ undergoes no structural transformation as a function of temperature, and the magnetic and charge order transitions are separated in temperature with the anti-ferromagnetic Neel temperature at $T_N \approx 120\,\mathrm{K}$ and charge disproportionation temperature at $T_{CD} \approx 343\,\mathrm{K}$ [21]. These observations are one motivation for the investigation of the local correlation driven charge disproportionation in Paper VI.

Part I

Models

Chapter 2

Hubbard Model

In 1963 Hubbard published a paper [11] where he deduced and motivated one of the simplest conceivable model Hamiltonians for interacting electrons in narrow bands. This model is today called the *Hubbard model* and has been extremely important for our understanding of correlated electron phenomena. Starting out from the very general model Hamiltonian in Eq. (1.10) we here summarize the approximation steps necessary to obtain the Hubbard model.

Assuming a narrow band implies that the Wannier orbitals $\psi_{\mathbf{R}n\sigma}(\mathbf{r})$ are strongly localized around the lattice sites $\mathbf{R}_i$. The resulting small overlap between Wannier functions on different sites causes the interaction integrals $V_{\alpha\beta\gamma\delta}$ in Eq. (1.9) to have a substantial value only on the same site

$$\begin{aligned} V_{\alpha\beta\gamma\delta} &= (\mathbf{R}_i n\sigma, \mathbf{R}_j m\sigma' | r_{12}^{-1} | \mathbf{R}_k n'\sigma, \mathbf{R}_l m'\sigma') \\ &\approx \delta_{\mathbf{R}_i\mathbf{R}_j}\delta_{\mathbf{R}_j\mathbf{R}_k}\delta_{\mathbf{R}_k\mathbf{R}_l}(\mathbf{R}_i n\sigma, \mathbf{R}_i m\sigma' | r_{12}^{-1} | \mathbf{R}_i n'\sigma, \mathbf{R}_i m'\sigma') \,. \end{aligned} \quad (2.1)$$

The localized Wannier states also gives small kinetic overlaps between next-nearest-neighbors so the hopping $t_{\mathbf{R}\mathbf{R}',n\sigma}$ can be approximately restricted to only nearest-neighbors

$$t_{\mathbf{R}\mathbf{R}',n\sigma} \approx \delta_{\langle \mathbf{R}\mathbf{R}' \rangle} t_{n\sigma} \,, \quad (2.2)$$

where $\langle \mathbf{R}\mathbf{R}' \rangle$ denotes nearest-neighbor sites. This restriction can in reality only be motivated for materials with bands having d- and f-character, but in the name of model simplicity we consider the equivalent system with a band having s-character. Describing a s-band requires only one Wannier function

so the band index n can be dropped. The model parameters then reduce to

$$\begin{aligned} V_{\alpha\beta\gamma\delta} &= \delta_{\mathbf{R}_i\mathbf{R}_j}\delta_{\mathbf{R}_j\mathbf{R}_k}\delta_{\mathbf{R}_k\mathbf{R}_l}\langle \mathbf{R}_i\sigma, \mathbf{R}_i\sigma' | r_{12}^{-1} | \mathbf{R}_i\sigma, \mathbf{R}_i\sigma'\rangle = U\,, \\ t_{\mathbf{R}\mathbf{R}',\sigma} &= -\delta_{\langle \mathbf{R}\mathbf{R}'\rangle} t\,, \end{aligned} \tag{2.3}$$

where the hopping also is assumed to be spin independent. The involved Hamiltonian in Eq. (1.10) for interacting electrons in a crystal has now been reduced to a model determined by a single parameter, namely the ratio U/t. Where the nearest-neighbor single particle hopping t describes the itinerant property of the electrons, and the extra energy contribution U when two electrons occupy the same lattice site approximate the electron-electron interaction. These approximations gives rise to the sought effective Hamiltonian, the *Hubbard Model*

$$\hat{H}_{\text{eff}} \approx \hat{H}_{\text{Hub}} = -t \sum_{\langle \mathbf{R}\mathbf{R}'\rangle} \sum_{\sigma} c^{\dagger}_{\mathbf{R}\sigma} c_{\mathbf{R}'\sigma} + U \sum_{\mathbf{R}} \hat{n}_{\mathbf{R}\uparrow}\hat{n}_{\mathbf{R}\downarrow}\,, \tag{2.4}$$

where $\hat{n}_{\mathbf{R}\sigma} = c^{\dagger}_{\mathbf{R}\sigma} c_{\mathbf{R}\sigma}$ is the electron number operator. One of the open questions is what sort of ordered phases can be found in the Hubbard model. At half-filling and at finite coupling $U \neq 0$, it is well established that the ground state is anti-ferromagnetic. But at finite doping there may be superconducting as well as spin and charge density wave states [22].

However our current focus is the Mott metal-insulator transition at half-filling. To this end we start out with discussing the basic thermodynamics of the Hubbard model in Section 2.1, then after a short introduction of the Green's function formalism in Section 2.2, we go on to discuss the non-interacting and atomic limits of the model in Section 2.3. We then treat the limit of infinite coordination number (or infinite dimension) in Section 2.4 and discuss the qualitative features of the finite temperature phase diagram at half-filling in Section 2.5. The chapter is closed by Section 2.6 that gives an introduction to the Landau theory phenomenological description of first order transitions.

2.1 Thermodynamics

As we are going to work with the metal-insulator phase transition of the Hubbard model let us first state some useful thermodynamic relations. It is convenient to work in the grand canonical ensemble, with variable particle number and fixed chemical potential μ. The grand partition function $\mathcal{Z}$ is

then given by

$$\mathcal{Z} = \mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}]\,, \tag{2.5}$$

where β is the inverse temperature, $\beta = (k_B T)^{-1}$, the trace runs over a complete set of states, and $\hat{N}$ is the total electron number operator, $\hat{N} = \sum_{i\sigma} \hat{n}_{i\sigma}$. In this formulation the expectation value of an operator $\hat{O}$ is defined through its mutual trace with the Gibbs distribution $\hat{\rho} = e^{-\beta(\hat{H}-\mu\hat{N})}/\mathcal{Z}$ [23]

$$\langle \hat{O} \rangle = \mathrm{Tr}[\hat{O}\hat{\rho}] = \mathcal{Z}^{-1}\mathrm{Tr}[\hat{O}e^{-\beta(\hat{H}-\mu\hat{N})}]\,. \tag{2.6}$$

Returning to the Hubbard model it can be divided in a kinetic part $\hat{T}$ and an interaction part $\hat{V}$ as follows

$$\begin{aligned}\hat{H} &= -t \sum_{\langle \mathbf{R}\mathbf{R}' \rangle} \sum_{\sigma} c^{\dagger}_{\mathbf{R}\sigma} c_{\mathbf{R}'\sigma} + U \sum_{\mathbf{R}} \hat{n}_{\mathbf{R}\uparrow}\hat{n}_{\mathbf{R}\downarrow} = \\ &= \hat{T} + U\hat{D} = \hat{T} + \hat{V}\,,\end{aligned} \tag{2.7}$$

where we also have introduced the total double occupancy operator, $\hat{D} = \sum_{\mathbf{R}} \hat{n}_{\mathbf{R}\uparrow}\hat{n}_{\mathbf{R}\downarrow}$. In Eq. (2.7) the Hubbard coupling U now enters the Hamiltonian as an external field coupled to the conjugate field $\hat{D}$ [24].

When studying first-order phase transitions one sometimes can find several coexisting phases, where the thermodynamical ground state is naturally the phase with the lowest free energy Ω, related to the Grand partition function $\mathcal{Z}$ through

$$\Omega = -\frac{1}{\beta}\ln \mathcal{Z}\,, \quad \mathcal{Z} = e^{-\beta\Omega}\,. \tag{2.8}$$

Taking the derivative of Ω with respect to U and using Eq. (2.7) gives

$$\begin{aligned}\frac{d\Omega}{dU} &= -\frac{1}{\beta}\frac{d}{dU}\ln \mathcal{Z} = -\frac{1}{\beta}\frac{1}{\mathcal{Z}}\frac{d}{dU}\mathrm{Tr}\left[e^{-\beta(\hat{H}-\mu\hat{N})}\right] = \\ &= \frac{1}{\mathcal{Z}}\mathrm{Tr}\left[e^{-\beta(\hat{H}-\mu\hat{N})}\frac{\partial}{\partial U}(\hat{H}-\mu\hat{N})\right] = \\ &= \frac{1}{\mathcal{Z}}\mathrm{Tr}\left[e^{-\beta(\hat{H}-\mu\hat{N})}\left(\hat{D} + \frac{\partial\mu}{\partial U}\hat{N}\right)\right] = \\ &= \langle \hat{D} \rangle + \frac{\partial\mu}{\partial U}\langle \hat{N} \rangle\,.\end{aligned} \tag{2.9}$$

Using this, the free energy Ω of the interacting system and the free energy Ω_0 of the non-interacting system can be connected through the integral

$$\Omega - \Omega_0 = \int_0^U d\tilde{U} \left(\langle \hat{D} \rangle + \frac{\partial \mu}{\partial U} \langle \hat{N} \rangle \right) , \tag{2.10}$$

and the free energy difference $\Delta\Omega$ between two systems with different couplings U_1 and U_2 can be calculated as

$$\Delta\Omega = \Omega_2 - \Omega_1 = \int_{U_1}^{U_2} d\tilde{U} \left(\langle \hat{D} \rangle + \frac{\partial \mu}{\partial U} \langle \hat{N} \rangle \right) . \tag{2.11}$$

Within a mean-field theory with a hysteresis region of coexisting solutions this can be used to locate the accompanied thermodynamical first-order transition [25]. By integrating along the adiabatic connection between coexisting solutions the transition occur at the coupling U where the coexisting solutions has zero free energy difference, $\Delta\Omega = 0$.

In preparation for the subsequent discussions let us end this thermodynamical discussion by defining the thermodynamical intensive variables corresponding to the extensive expectation-values $\langle \hat{N} \rangle$ and $\langle \hat{D} \rangle$

$$N = \frac{\langle \hat{N} \rangle}{\mathcal{N}} , \quad D = \frac{\langle \hat{D} \rangle}{\mathcal{N}} , \tag{2.12}$$

where $\mathcal{N}$ is the total number of lattice sites and N and D are the average number of electrons and double occupancies per lattice site.

2.2 Matsubara Green's functions

The success of DFT mentioned in the introduction stems from its use of a non-interacting reference system. Although algorithmically efficient, this approach become increasingly problematic as the local electron-electron interaction effects become prominent. The language of choice for describing systems with many-body interactions are the Green's function formalism [26, 27].

The single particle Green's function G describes the propagation of an electron with spin σ added to the system at a position $\mathbf{R}'$ and time t' that is later removed from the system at $\mathbf{R}$, t

$$G_\sigma(\mathbf{R}, t; \mathbf{R}', t') = -i \langle c_{\mathbf{R}\sigma}(t) c^\dagger_{\mathbf{R}'\sigma}(t') \rangle . \tag{2.13}$$

For finite-temperature systems the inverse temperature β enters in the same way as time but with an imaginary prefactor and can be described using a Wicks-rotation of time t to imaginary time τ. Further the τ dependence of operators are defined in the Heisenberg representation [14] as

$$\begin{aligned} c^\dagger_{\mathbf{R}\sigma}(\tau) &= e^{\tau(\hat{H}-\mu\hat{N})} c^\dagger_{\mathbf{R}\sigma} e^{-\tau(\hat{H}-\mu\hat{N})}\,, \\ c_{\mathbf{R}\sigma}(\tau) &= e^{\tau(\hat{H}-\mu\hat{N})} c_{\mathbf{R}\sigma} e^{-\tau(\hat{H}-\mu\hat{N})}\,. \end{aligned} \tag{2.14}$$

The real-space $\mathbf{R}$ and imaginary time τ dependence of the Green's function can be transformed into momentum $\mathbf{k}$ and Matsubara frequency, $\omega_n = \frac{\pi}{\beta}(2n+1)$ dependence, which simplifies calculations for translational invariant time independent systems. The single-particle Green's function G_0 for a free system has the form

$$G_0(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n - \epsilon_{\mathbf{k}} + \mu}\,, \tag{2.15}$$

where $\epsilon_{\mathbf{k}}$ is the single particle dispersion. The power of using single-particle Green's functions is that all contributions from interactions enter the interacting Green's function G only through the self-energy $\Sigma(\mathbf{k}, i\omega_n)$

$$G(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n - \epsilon_{\mathbf{k}} + \mu - \Sigma(\mathbf{k}, i\omega_n)}\,. \tag{2.16}$$

In the language of Feynman diagrams the self-energy Σ is given by the sum of all one-particle irreducible amputated diagrams connecting to two external propagators [14]. In general it is not possible to calculate Σ using this definition, but it can be used to formulate approximations in the weak coupling limit by truncating the sum. To go to higher orders is increasingly difficult as both the number of diagrams and the number of required integrations over the Brillouin zone increase dramatically.

This section should not be seen as a sufficient introduction Green's function theory and its purpose is only to establish the notation. The reader who wishes to get a throughout treatment of the subject is referred to one of the works [14, 26, 27]. Now we will take the presented Green's function formalism and apply it to the Hubbard model in some simple limits.

2.3 Non-interacting and atomic limits

Before discussing the phase diagram of the Hubbard model it is fruitful to study more in detail some of the extreme limits. In particular the non-

interacting and atomic limits are of interest. With these limits as a starting-point we will try to make it plausible that there should exist an intermediate phase transition.

In the *non-interacting limit* the electron-electron interaction is removed by letting $U = 0$, and the system simplifies to a single band of non-interacting electrons with nearest neighbor hopping only. As in Section 1.1, the single particle hopping is diagonal in momentum space, enabling us to rewrite the Hamiltonian as

$$\hat{H}_{U=0} = -t \sum_{\langle \mathbf{R}\mathbf{R}' \rangle} \sum_{\sigma} c^{\dagger}_{\mathbf{R}\sigma} c_{\mathbf{R}'\sigma} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} \,, \tag{2.17}$$

where $\epsilon_{\mathbf{k}}$ is the dispersion of the electrons given by

$$\epsilon_{\mathbf{k}} = -t \sum_{\langle \mathbf{R}, 0 \rangle} e^{-i\mathbf{k}\cdot\mathbf{R}} \,, \tag{2.18}$$

from which the density of states (DOS) is directly obtained as

$$\rho(\epsilon) = \sum_{\mathbf{k}} \delta(\epsilon - \epsilon_{\mathbf{k}}) \,, \tag{2.19}$$

and the Green's function is given by

$$G(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n + \mu - \epsilon_{\mathbf{k}}} \,. \tag{2.20}$$

The system is metallic whenever the DOS is non-zero at the chemical potential, $\rho(\mu) \neq 0$. Translated to the number of electrons per site N

$$N = \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) f(\epsilon) \,, \tag{2.21}$$

where $f(\epsilon) = (e^{\beta(\epsilon-\mu)} + 1)^{-1}$ is the Fermi distribution function, we see that the system is metallic whenever the band is not completely filled, $N < 2$. Some lattice specific density of states will be shown in next section. The on-site double occupancy D can be calculated using Wick's theorem [14]

$$D = \frac{\langle \hat{D} \rangle}{\mathcal{N}} = \langle \hat{n}_{\mathbf{R}\uparrow} \hat{n}_{\mathbf{R}\downarrow} \rangle = \langle \hat{n}_{\mathbf{R}\uparrow} \rangle \langle \hat{n}_{\mathbf{R}\downarrow} \rangle \tag{2.22}$$

which for the half-filled paramagnetic system gives, $D = 1/4$.

The other extreme case is *the atomic limit* where the bandwidth goes to zero and the electron dispersion $\epsilon_{\mathbf{k}}$ becomes $\mathbf{k}$ independent. This makes the single-particle "hopping" local by

$$\epsilon_{\mathbf{k}} = t_0 \quad \Rightarrow \quad t_{\mathbf{R}\mathbf{R}'} = \delta_{\mathbf{R}\mathbf{R}'} t_0 \,, \tag{2.23}$$

resulting in a Hamiltonian that is diagonal in real space

$$\hat{H}_{\text{atm}} = \sum_{\mathbf{R}\sigma} t_0 c^\dagger_{\mathbf{R}\sigma} c_{\mathbf{R}\sigma} + U \sum_{\mathbf{R}} c^\dagger_{\mathbf{R}\uparrow} c_{\mathbf{R}\uparrow} c^\dagger_{\mathbf{R}\downarrow} c_{\mathbf{R}\downarrow} \,. \tag{2.24}$$

The local Hamiltonian $\hat{H}_{\text{atm}}$ is also diagonal in the occupation number basis, $|0\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, $|\uparrow\downarrow\rangle$, and its Green's function G_{atm} can be calculated using the Lehman spectral representation [14] giving

$$G_{\text{atm},\sigma}(i\omega_n) = \frac{1 - \langle \hat{n}_{\bar{\sigma}} \rangle}{i\omega_n - t_0 + \mu} + \frac{\langle \hat{n}_{\bar{\sigma}} \rangle}{i\omega_n - t_0 - U + \mu} \,, \tag{2.25}$$

where $\bar{\sigma}$ denotes the opposite spin z-component of σ. Rewriting the Green's function on the form of Eq. (2.16) the self-energy contribution can readily be determined to

$$\Sigma_{\text{atm},\sigma}(i\omega_n) = \langle \hat{n}_{\bar{\sigma}} \rangle U + \frac{\langle \hat{n}_{\bar{\sigma}} \rangle (1 - \langle \hat{n}_{\bar{\sigma}} \rangle) U^2}{i\omega_n - t_0 + \mu - (1 - \langle \hat{n}_{\bar{\sigma}} \rangle) U} \,. \tag{2.26}$$

The density of states is also momentum independent having the form

$$\rho_\sigma(\epsilon) = (1 - \langle \hat{n}_{\bar{\sigma}} \rangle) \delta(\epsilon - t_0 + \mu) + \langle \hat{n}_{\bar{\sigma}} \rangle \delta(\epsilon - t_0 - U + \mu) \,, \tag{2.27}$$

with two resonance peaks separated by U, giving an insulator at half-filling. As the occupied low energy resonance only have contributions from singly occupied states the double occupancy is zero at half-filling

$$D = \langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle = 0 \,. \tag{2.28}$$

Now we can compare the two extreme cases, the non-interacting and the atomic limit and their respective density of states, see Fig. 2.1. If we treat U/t as a tunable parameter we see that the system goes from a metallic state in the non-interacting limit, $U/t = 0$, to an insulating state in the atomic-limit, $U/t = \infty$. Thus at some intermediate value of the electron-electron interaction U one should expect some form of phase transition going from a metallic to an insulating density of states. This interaction driven transition was first proposed by Nevill F. Mott in 1949 and is therefore called the Mott-transition [28]. The transition occurs approximately when the coupling U is of the same order as the bandwidth, $W \propto t$, of the non-interacting system, $U/W \sim 1$. For this reason it is common to use the ratio U/W instead of U/t to specify the free parameter in the Hubbard model.

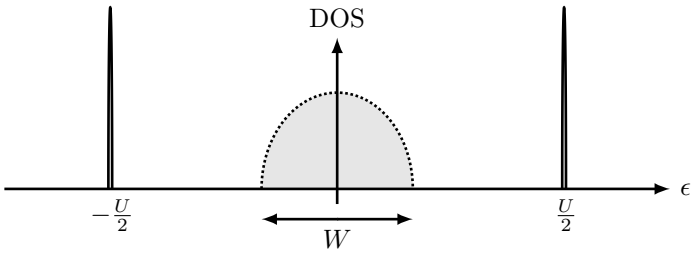


Figure 2.1: Schematic density of states in the metallic non-interacting limit (dotted line and gray shaded area) with a bandwidth of W and insulating atomic-limit (solid lines).

2.4 Limit of infinite dimensions

While the atomic and non-interacting limits in the Hubbard model are important they do not contain any interesting many-body effects. Another limit that actually does, is the limit of infinite dimensions, which corresponds to the case when the number of nearest neighbors – i.e. the coordination number z – goes to infinity. The prescription on how to take this limit was presented in the seminal paper of Metzner and Vollhardt [29]. This summary follows the appendix on Fermiology in Ref. [17].

Consider a hyper-cubic lattice in d dimensions with nearest neighbor hopping, and a lattice spacing set to one. Then the dispersion in Eq. (2.18) takes the form

$$\epsilon_{\mathbf{k}} = -2t \sum_{i=1}^{d} \cos k_i \,. \tag{2.29}$$

When taking the limit $d \to \infty$ the sum over the $\cos k_i$ terms for any $\mathbf{k}$ becomes essentially a sum over independent random values each with a cosine distribution. The central limit theorem can then be applied to obtain the distribution of ϵ that we usually call the density of states $\rho(\epsilon)$ in the form of a normal distribution

$$\rho(\epsilon) = \frac{1}{2t\sqrt{\pi d}} e^{-\left(\frac{\epsilon}{2t\sqrt{d}}\right)^2} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\epsilon^2}{2\sigma^2}} \,, \tag{2.30}$$

with a variance σ^2 depending on t as $\sigma = \sqrt{2d}\,t$. In order to get a finite density of states when $d \to \infty$, the hopping has to be rescaled as

$$t \to \frac{t}{\sqrt{2d}} = \frac{t}{\sqrt{z}} \,, \tag{2.31}$$

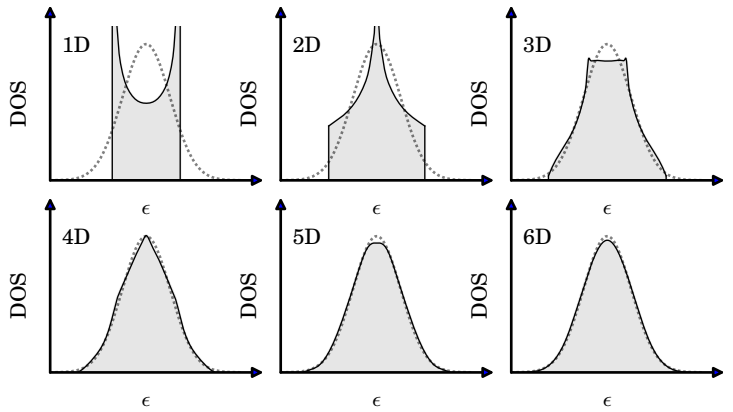


Figure 2.2: Density of states for the square lattice in 1 to 6 dimensions (solid lines and gray shaded areas) compared with the infinite dimensional limit (gray dotted line), inspired by [17].

where we in the last step have used that the coordination number z of the hyper-cubic lattice is two per dimension, $z = 2d$. This scaling results in non-trivial many-body physic for the Hubbard model even in the limit of $d, z \to \infty$ [17].

An interesting question is how close the limit of infinite coordination number real crystalline materials actually are. For example the three dimensional lattices, simple cubic, body centered cubic and face centered cubic has coordination numbers, $z = 6$, 8 and 12 respectively. The DOS of the hyper-cubic lattice is shown in Fig. 2.2, for dimensions 1 to 6 and ∞, i.e. for even coordination numbers $z = 2$ to 12 and ∞. Studying the evolution of the DOS as a function of dimension it is evident that already in four dimensions and at coordination number $z = 8$, almost all features of the finite dimensional lattice is washed out.

While the hypercubic lattice is a nice example of the infinite dimensional limit its infinite band width makes it qualitative different from finite dimensional systems. Most of the model calculations in this thesis has instead been performed for the Bethe lattice, with finite bandwidth. It is one of the most popular reference systems with large amounts of available reference results. Strictly speaking the Bethe lattice is not a lattice but a graph with fixed coordination number z, having the special property that if one site is removed the graph is divided into z disconnected pieces. It is also a bipartite graph as the sites can be partitioned in two groups, A and B having

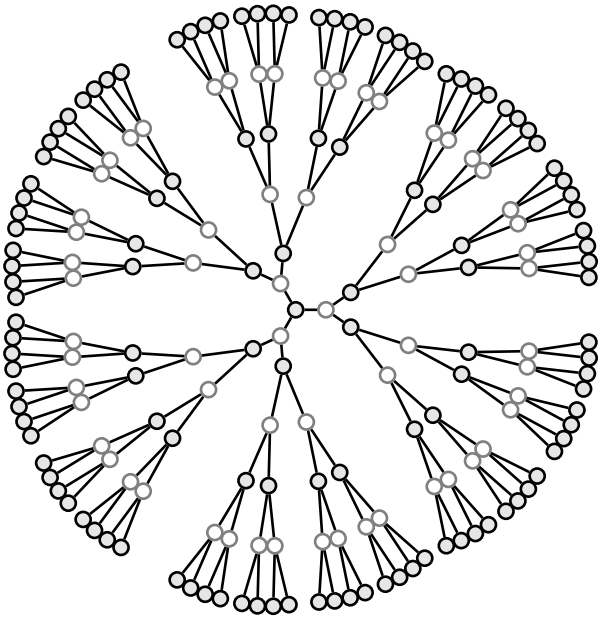


Figure 2.3: Part of the Bethe lattice with coordination number, $z = 3$, with all nearest, next-nearest up to (next)^5-nearest neighbors of the central site. The bipartite property is shown by the color coding of the sites for the A (black circles) and B (gray circles) sub-lattices.

nearest neighbors of the other group only. A sketch of a Bethe lattice with coordination $z = 3$ is shown in Fig. 2.3.

In the limit of infinite coordination number the non-interacting DOS $\rho^{(0)}(\epsilon)$ of the Bethe lattice becomes semi-circular, with a bandwidth $W = 4t$

$$\rho^{(0)}(\omega) = \frac{2}{\pi}\sqrt{1 - \left(\frac{2\omega}{W}\right)^2}, \quad |\omega| < \frac{W}{2}\,. \tag{2.32}$$

The finite bandwidth of the Bethe lattice is an important property when studying the Mott transition in the Hubbard model and the special topology simplifies to some extent the derivation of the Dynamical Mean Field self-consistency equations, see Chapter 4.

2.5 Phase diagram

As mentioned above the single parameter determining the Hubbard model is the ratio U/W (or U/t), but when studying its phase diagram the other external parameters that must be considered are, the temperature T (or the inverse temperature $\beta = (k_B T)^{-1}$), and the average electron occupation N. At this point a general account for the structure of the phase diagram would be in place, but in this the particular lattice structure play a decisive role. Still it is possible to draw some general conclusions for bipartite lattices with coordination numbers larger than 4, a category containing for example the cubic lattice in 2 and 3 dimensions.

First limiting the discussion to zero temperature $T = 0$, the non-interacting system ($U = 0$) is paramagnetic (PM) for all fillings $0 \leq N \leq 2$. While at half-filling $N = 1$, an infinitesimal U is sufficient to make the ground state anti-ferromagnetic (AF). For larger values of U this AF region persist for larger deviations of N away from half-filling. Lastly there are studies indicating that there exists a ferromagnetic state (FM) for very high values of U, a result that is readily obtained in mean-field approximation methods. There seems to still be a controversy whether this is a result of the approximate methods or a real phase of the Hubbard model [22].

For finite temperatures, the paramagnetic phase persist, while in the N and U regions of the zero-temperature AF ground state, a transition to the PM state occur at some Neel temperature T_N, shown schematically for half-filling in Fig. 2.4a. The high-temperature PM state at half-filling is divided in two regions going from a good metal at low U to an insulator when increasing U, with a crossover region separating the phases [30, 31].

This tendency towards AF is physically a pathology of the over-simplified Hubbard model on bipartite lattices with only nearest-neighbor hopping. The fermi surface has perfect nesting and any perturbation drives the instability towards an AF state. In more complicated models and real correlated electron-systems the anti-ferromagnetism is often suppressed by next-nearest neighbor hopping or inter-band electron-electron interaction. Thus, to connect results for the simple Hubbard model to these classes of models the AF state is often removed by hand in calculations.

Now limiting the discussion to half-filling, $N = 1$, and Fig. 2.4, we superficially suppress the anti-ferromagnetic state restricting the system to be paramagnetic. This reveals a transition line between the metal and insulator as anticipated in previous section. This first-order transition line ends in a critical end-point at a coupling U_c and temperature T_c, that is also the locus of the high temperature crossover region, see Fig 2.4b. For lattices

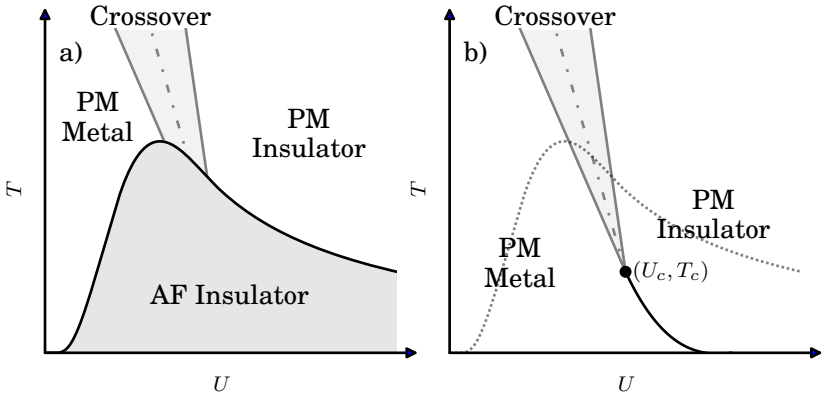


Figure 2.4: Qualitative phase diagram of the Hubbard model at half-filling on a bipartite lattice, a) real phases with the low-temperature anti-ferromagnetic (AF) ground state and high temperature paramagnetic (PM) metallic and insulating phases separated by a crossover region, b) phase diagram with suppressed AF showing the first-order metal insulator transition line (solid black line) ending in the second-order critical endpoint at (U_c, T_c) (black dot). The asymptotic behaviors of the Neel temperature are taken from [31].

with non-interacting DOS having bandwidths W of the order of one electron volt, $W \sim 1\,\text{eV}$, the critical temperature T_c has the same order of magnitude as room temperature $T_c \sim 25\,\text{meV}$. The properties of this transition are studied in detail in Paper I where we determine the critical exponents of the end point of the paramagnetic first order transition.

2.6 Landau theory

To assist in the understanding of the finite temperature critical point in the Hubbard model treated by Paper I, a brief introduction to Landau Theory is in place. In the paper this theoretical framework is used to analyze the critical behavior.

Landau Theory [32] is a phenomenological theory aimed at describing phase transitions in thermodynamical systems. First the state of the system is characterized by some thermodynamical internal observable, conjugate to the external fields, and its behavior as a function of temperature T and the external fields are considered. In general whenever the system undergo a first-order transition the conjugate variable will develop non-analytic be-

havior. This is modeled in Landau Theory by introducing a continuous Landau function $\mathcal{L}$, parametrized by the observable, the external fields and temperature, whose global minimum reflects the state of the system. In this way the non-analytic behavior of first-order transitions can be described by jumps between local minima of $\mathcal{L}$. This section is an adaptation of the ideas in Negele and Orland [14] to the Mott-transition in the Hubbard model.

For the Hubbard model we have already seen in Eq. (2.7) that the external field, relevant for the Mott metal-insulator transition, is the Hubbard coupling U which together with the temperature T determines the external parameters of the model. The idea of using the on-site double occupancy D as the variable conjugate to U was already mentioned in Section 2.1, and can be traced back to Castellani [24]. Thus one can try to describe the Mott-transition by introducing a continuous function $\mathcal{L}(U, T, D)$ whose global minima in D coincides with the state of the Hubbard model. As the Mott-transition has a second-order critical end-point at the critical coupling and temperature (U_c, T_c), we can try to expand $\mathcal{L}$ in powers of D around this point. In terms of the shifted variables

$$u = U - U_c\,, \quad t = T - T_c\,, \quad d = D - D_c\,, \tag{2.33}$$

the expansion to fourth order has the form

$$\mathcal{L}(u, t, d) = a_0(u, t)d + a_1(u, t)\frac{d^2}{2} + a_2(u, t)\frac{d^3}{3} + \frac{d^4}{4}\,, \tag{2.34}$$

where the coefficients $a_n(u, t)$ can be taken to linear order in u and t, $a_n(u, t) = \beta_n u + \gamma_n t$. It is not directly evident that this expansion can describe the physics of the transition in the vicinity of the critical point, but this question is investigated in Paper I. The reason for limiting the expansion to fourth order is that it is the lowest power expansion that can describe a second-order transition in the immediate vicinity of a first-order transition. When constructing the power expansion of the Landau function one can often reduce the number of parameters γ_n and β_n by imposing the symmetries of the system on $\mathcal{L}$. Unfortunately this is not the case for the Mott metal-insulator transition. Still some intuitive understanding of the expansion in Eq. (2.34) can be gained by studying a simple special case.

Consider the expansion

$$\mathcal{L}(u, t, d) = \beta_0 u d + \gamma_1 t\frac{d^2}{2} + \frac{d^4}{4}\,, \tag{2.35}$$

where the coupling u only enters in the linear term, the temperature t in the quadratic term and the coefficients are positive, $\beta_0 > 0$, $\gamma_1 > 0$. The state

of the system for fixed external parameters u and t is now determined by the global minima of $\mathcal{L}$ with respect to d. Plotting this minima in (u, t, d) space generates the surface $S_{\min[\mathcal{L}]}$

$$S_{\min[\mathcal{L}]} = \{(u, t, d) \,:\, \min_d[\mathcal{L}(u, t, d)]\}\,, \tag{2.36}$$

shown in Fig. 2.5. For $u = 0$ and $t < 0$ the system undergoes a first-order transition as can be inferred from the discontinuity of $S_{\min[\mathcal{L}]}$. The evolution into a first-order transition through a second-order critical point can be understood in terms of the d dependence of $S_{\min[\mathcal{L}]}$ varying u along an isotherm. In the right back plane of Fig. 2.5 the isotherms are shown, going from continuous behavior for $t > 0$, to a second-order transition at $u = 0$ and $t = 0$, which develops into a first-order transition for $t < 0$. Thus the phase diagram of the system is a first-order transition line ending in a second-order critical point, shown in the bottom plane of Fig. 2.5.

Describing the system with the continuous Landau function $\mathcal{L}(u, t, d)$ the non-analytic first-order behavior is generated through the competition between local minima. Above the critical temperature ($t > 0$) the system has a single global minima, and no phase transition occur when changing the external field u. At the critical point ($u = 0, t = 0$) the global minimum Landau function has zero first, second, and third order derivatives, thus yielding a second-order transition $\partial_d\mathcal{L}(0, 0, 0) = 0$, $\partial_d^2\mathcal{L}(0, 0, 0) = 0$, $\partial_d^3\mathcal{L}(0, 0, 0) = 0$. Below the critical temperature ($t < 0$) the Landau function has two local minima for small values of u. For $u \neq 0$ one of these minima is also the global minima, but when $u = 0$ both minima becomes global and the system jumps discontinuously from one to the other for any perturbations in u. The stationary points of $\mathcal{L}$ in the vicinity of the critical point are shown in Fig. 2.6.

In order to connect to the results in Paper I it is not sufficient to study the surface of system states given by the global minimum of $\mathcal{L}$. Instead one have to consider all stationary points of $\mathcal{L}$ and investigate the surface of stationary points $S_{\mathcal{L}}$ defined by

$$\begin{aligned} S_{\mathcal{L}} =& \{(u, t, d) \,:\, \partial_d\mathcal{L}(u, t, d) = 0\}\,, \\ \partial_d\mathcal{L} =& \beta u + \gamma t d + d^3 = 0\,. \end{aligned} \tag{2.37}$$

In the system state space $S_{\mathcal{L}}$ is a continuous smooth surface where the extra stationary points excluded from $S_{\min[\mathcal{L}]}$ now continuously connects the states at the discontinuous first-order phase transition, see Fig. 2.7. The topology of $S_{\mathcal{L}}$ is in catastrophe theory called *the cusp catastrophe* [33] being a surface in three dimensions with a single fold. A cusp catastrophe

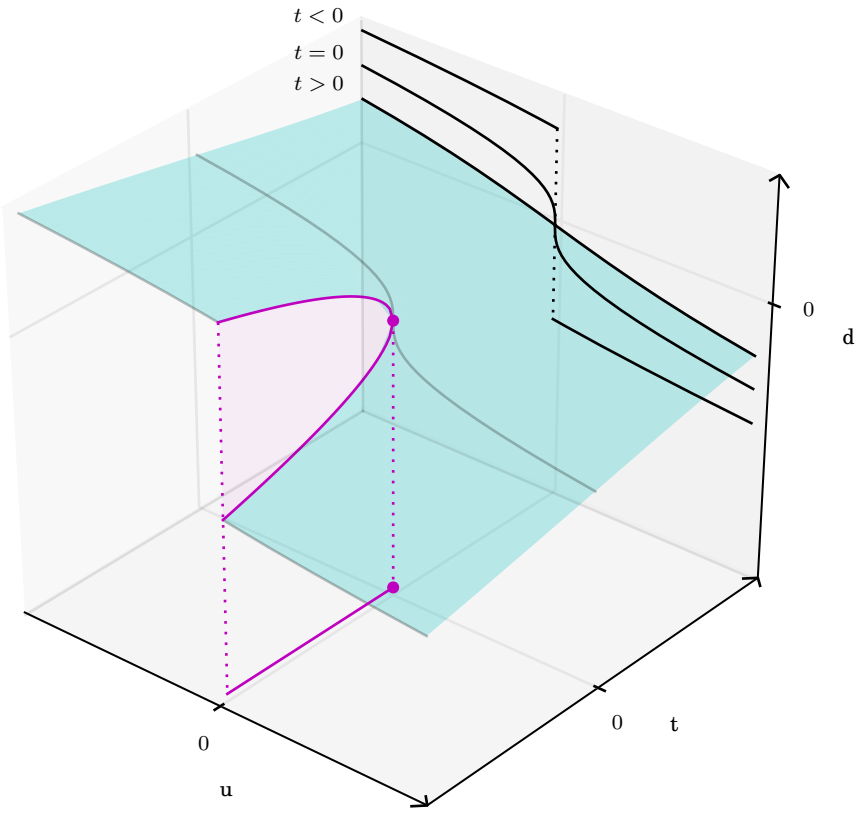


Figure 2.5: State of the system in (u, t, d) space. The first-order transition is accompanied with a discontinuity at $u = 0$ and $t < 0$ marked in magenta. The phase diagram in the (u, t) plane (bottom plane) has a first-order transition line (magenta) that terminates in a second-order critical end-point (magenta circle). The isotherms (right back-plane) are continuous for $t > 0$, singular at $t = 0$ and become discontinuous for $t < 0$.

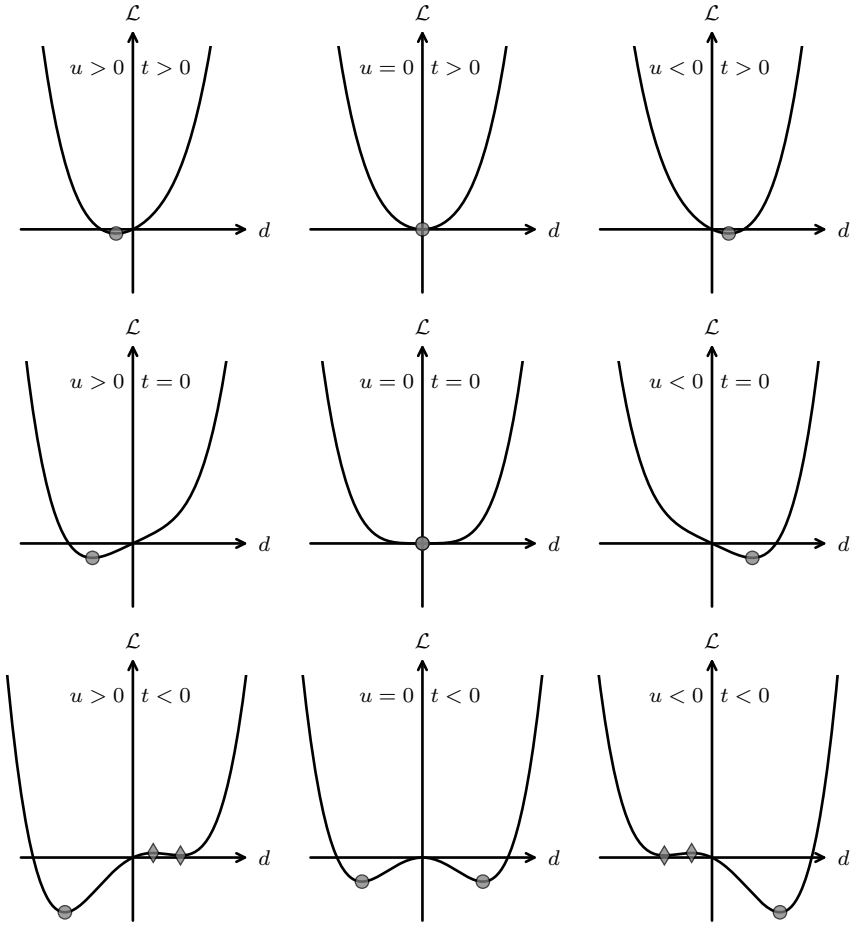


Figure 2.6: Behavior of the Landau function $\mathcal{L}$ at and around the second-order critical end-point. The global minima (circles) and the other stationary points (diamonds) are also shown. In the lowest center plot the first-order transition is show with its pair of separated global minima. (Inspired by [14])

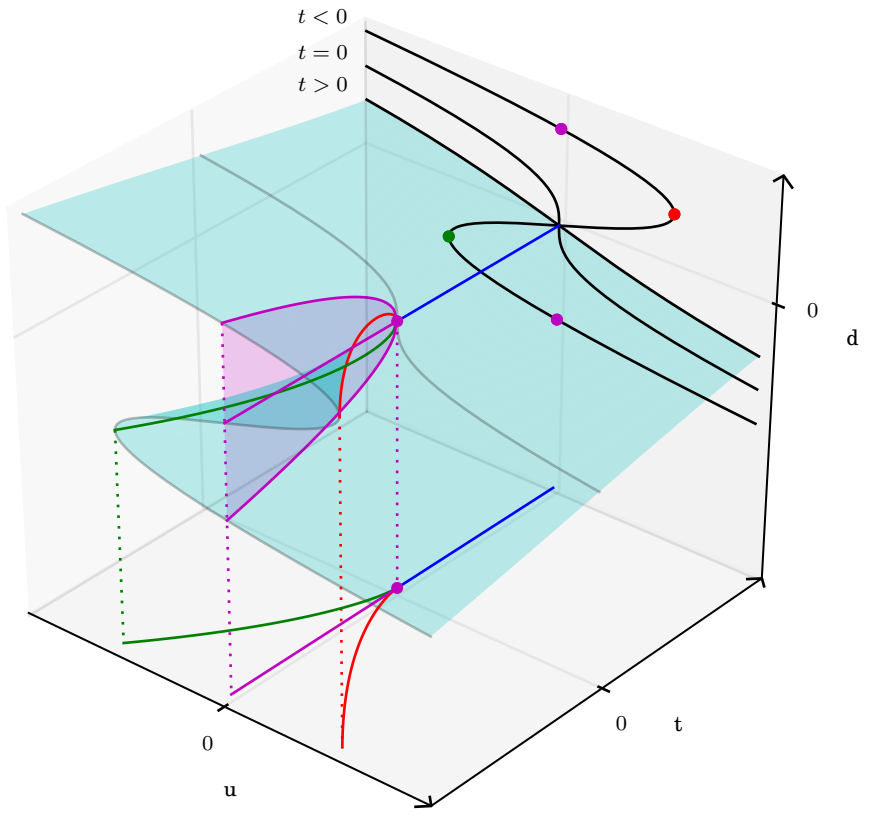


Figure 2.7: The continuous surface $S_{\mathcal{L}}$ of stationary points (cyan) of the Landau function $\mathcal{L}(u, t, d)$ and the cusp catastrophe at the critical point $(u, t, d) = (0, 0, 0)$. The first-order transition (magenta circles) on the isotherm ($t < 0$) (right back plane) is now continuously connected through a Z-shaped curve of stationary points combining through saddle-node bifurcations (red and green circle). The coexistence of stationary points now surrounds the first-order transition (magenta line in the bottom plane) with a hysteresis region (red and green lines) that mark the saddle-node bifurcation of two stationary points. Along the surface of the first-order transition (magenta) a pitchfork bifurcation occurs at the second-order critical point (central magenta circle).

is accompanied with a set of general properties, for example the first-order transition is surrounded by a hysteresis region in the (u, t) plane where three stationary points coexists, see the bottom plane in Fig. 2.7. Also on the first-order transition surface of $S_{\min[\mathcal{L}]}$ the stationary points exhibit a pitchfork bifurcation at the critical point, where a single stationary point at $t > 0$ separates in to two local minima and one intermediate local maximum. This type of pitchfork bifurcation is sometimes denoted as supercritical, see blue and magenta lines in Fig. 2.7. The boundaries of the hysteresis region are given as the point where one of the local minima annihilate with the local maximum in a tangent-bifurcation, as seen for the isotherm with $t < 0$ in the right back plane of Fig. 2.7. The maximum and minimum merge in to a saddle-point of $\mathcal{L}$ causing the second order derivative to be zero, $\partial_d^2 \mathcal{L} = 0$. Understanding the behavior of $S_{\mathcal{L}}$ and Fig. 2.7 greatly facilitate the interpretation of the results for the single band Hubbard model.

Chapter 3

Multi-band Models

To study correlated transition metal compounds we return to the full electron problem in Eq. (1.2), and try to isolate the relevant low energy degrees of freedom. In these systems correlated phenomena arise when the narrow transition metal d-band is situated close to the Fermi level.[1] Furthermore, it is in the vicinity of the corresponding Fermi surfaces that low energy excitations can occur, and these excitations are directly affected by local interactions. So a minimal requirement on a low energy effective model is a correct description of the Fermi surface structure.

Returning to the single-band Hubbard model of Chapter 2, its single band give rise to a single Fermi surface. This situation seldom occurs for transition metal compounds containing ten d-bands, that in general give rise to several fermi surfaces. Thus, even though the single-band Hubbard model exhibit most of the basic correlation phenomena, its single Fermi surface prohibits a quantitative description of most transition metal compounds. The remedy for this situation is the multi-band generalization of the Hubbard model.

The chapter is organized as follows: Section 3.1 describes the Projected Local Orbital (PLO) scheme as one possible approach for constructing local single-particle Wannier functions with d-electron angular momentum symmetry. In this basis the general local multi-band interaction is readily written down in Section 3.2. As the interaction commonly is assumed to be rotationally invariant, Section 3.3 investigates the most general form of two-particle interaction operators with this property. We then become more specific and derive the general form of the local Coulomb interaction in Section

[1] On the DFT level of description.

3.4 assuming rotational invariance, which gives the so called Slater-Condon form of the Coulomb interaction. The atomic problem is discussed in Section 3.5 by the introduction of quasi-spin that together with total particle number, spin, and angular momentum gives a complete many-body basis for the local d-electron Fock-space. The Kanamori interaction is introduced in Section 3.6, and in Section 3.7 we show that it is an approximation of the Slater-Condon interaction when applied to the entire set of d-electrons. This opens up for the possibility to parametrize the Slater-Condon interaction in terms of the Kanamori parameters as outlined in Section 3.8. The chapter is closed by Section 3.9, where other uses of the Kanamori parameters in the literature are discussed.

3.1 Single-particle wave-functions

Consider once more a general electron model as in Eq. (1.3) described by the Bloch wave-functions $\psi_{\mathbf{k}n\sigma}(\mathbf{r})$ with dispersion $\epsilon_{\mathbf{k}n\sigma}$, and the corresponding local Wannier functions [Eq. (1.4)]

$$\psi_{\mathbf{R}n\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \psi_{\mathbf{k}n\sigma}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{R}}\,, \tag{3.1}$$

where $\mathbf{R}$ is a lattice vector, N is the number of lattice sites, and n and σ are the Bloch band and spin indices respectively. For this system the Fermi surface structure is described by the low energy subset $\mathcal{C}$ of Wannier functions $\psi_{\mathbf{R}n\sigma}$ that cross the Fermi level ϵ_F at some point in the Brillouin zone, $\mathcal{C} = \{(n,\sigma) : \exists \mathbf{k}, \epsilon_{\mathbf{k}n\sigma} = \epsilon_F\}$. This procedure unambiguously identifies the set of states $\mathcal{C}$ needed to build a low energy effective model with preserved Fermi surface structure.

Unfortunately the local Wannier wave-functions $\psi_{\mathbf{R}n\sigma}$ do not have well defined symmetry properties. Actually, there is a lot of freedom in the local Wannier function construction, not shown in Eq. (3.1). Any k-dependent unitary transform $\mathcal{U}_{an}(\mathbf{k})$ defines an alternative set of orthonormal Wannier wave-functions

$$\psi_{\mathbf{R}a\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \sum_{n} \mathcal{U}_{an}(\mathbf{k}) \psi_{\mathbf{k}n\sigma}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}}\,. \tag{3.2}$$

The freedom in choosing $\mathcal{U}$ can be exploited for different purposes. For example one can minimize the spatial extent of the Wannier functions $\psi_{\mathbf{R}a\sigma}(\mathbf{r})$ in real space, which yields so called Maximally Localized Wannier functions (MLWFs) [34].

When the transition metal d-band is well separated from other bands, the MLWF procedure gives local wave-functions with well defined angular momentum character $l = 2$. However when the d-band is strongly hybridized and overlapping with ligand p-bands the localization can produce Wannier functions of mixed p and d character. This is problematic because, as we will soon see, the modeling of the local interactions assumes states with known local angular momentum quantum numbers. One approach that preserves the angular momentum character of the Wannier functions is to simply use the first steps in the MLWF algorithm, and stop before minimizing the spatial spread of the wave-functions. This class of wave-functions are sometimes referred to as projected local orbital Wannier functions (PLOWFs) [35]. Here we give a brief description of their construction.

Following Ref. [35] the first step is to define a set of local wave-functions $\chi_{m\sigma}(\mathbf{r})$, with good total angular momentum l and z-component m quantum numbers. This means that $\chi_{m\sigma}(\mathbf{r})$ has the same angular dependence as the spherical harmonic functions $Y_m^l(\Omega)$, and a typical choice is to simply use the corresponding wave-functions of the isolated atom. A related set of Wannier functions can then be constructed as

$$\psi_{\mathbf{R}m\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \sum_{n} P^{\sigma}_{mn}(\mathbf{k}) \psi_{\mathbf{k}n\sigma}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{R}} \,, \tag{3.3}$$

where the unitary projection matrix P is the overlap between the Bloch and local wave-functions

$$P^{\sigma}_{mn}(\mathbf{k}) = (\chi_{m\sigma} | \psi_{\mathbf{k}n\sigma}) \,. \tag{3.4}$$

The unitarity of P is guaranteed by the completeness of the Bloch wave-functions, which no longer apply when the sum over n is limited to the subset $\mathcal{C}$. To counter this, P can be replaced by the re-orthonormalized projection matrix

$$\bar{P}^{\sigma}_{mn}(\mathbf{k}) = \sum_{m'} (\sqrt{\mathcal{O}^{\sigma}(\mathbf{k})})^{-1}_{mm'} P^{\sigma}_{m'n}(\mathbf{k}) \,, \tag{3.5}$$

where $\mathcal{O}$ is the overlap matrix

$$\mathcal{O}^{\sigma}_{mm'}(\mathbf{k}) = \sum_{n \in \mathcal{C}} (P^{\sigma}_{mn}(\mathbf{k}))^* P^{\sigma}_{m'n}(\mathbf{k}) \,. \tag{3.6}$$

This gives a set of orthonormal Wannier wave-functions for the low energy subspace as

$$\psi_{\mathbf{R}m\sigma}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \sum_{n \in \mathcal{C}} \bar{P}^{\sigma}_{mn}(\mathbf{k}) \psi_{\mathbf{k}n\sigma}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}} \,, \tag{3.7}$$

where $\psi_{\mathbf{R}m\sigma}$ approximately retains the spherical harmonic angular dependence $\psi_{\mathbf{R}m\sigma}(\mathbf{r}) \propto Y_m^l(\Omega)$. Here on we will take the angular properties of $\psi_{\mathbf{R}m\sigma}$ for granted, but keep in mind that the accuracy of this approximation depends on the particular system and on the extent of the subset $\mathcal{C}$ (see also the discussion in Ref. [35]).

Passing over to second-quantization, let $c^\dagger_{\mathbf{R}m\sigma}$ denote the fermion creation operator for the state $\psi_{\mathbf{R}m\sigma}$ at lattice site $\mathbf{R}$ with angular and spin z-components m and σ respectively. The local total orbital angular momentum operators can then be defined as [36]

$$\begin{aligned}\hat{L}_z(\mathbf{R}) &= \sum_{m\sigma} m\, c^\dagger_{\mathbf{R}m\sigma} c_{\mathbf{R}m\sigma}\,, \\ \hat{L}_\pm(\mathbf{R}) &= \sum_{m\sigma} \sqrt{l(l+1) - m(m\pm 1)}\, c^\dagger_{\mathbf{R},m\pm1,\sigma} c_{\mathbf{R}m\sigma}\,,\end{aligned} \tag{3.8}$$

and similarly the local total spin operators take the form

$$\begin{aligned}\hat{S}_z(\mathbf{R}) &= \sum_{m\sigma} \sigma\, c^\dagger_{\mathbf{R}m\sigma} c_{\mathbf{R}m\sigma}\,, \\ \hat{S}_\pm(\mathbf{R}) &= \sum_{m\sigma} \sqrt{s(s+1) - \sigma(\sigma\pm 1)}\, c^\dagger_{\mathbf{R}m,\sigma\pm1} c_{\mathbf{R}m\sigma}\,.\end{aligned} \tag{3.9}$$

Translated to the Wannier basis, the kinetic Hamiltonian of Eq. (1.3) can be expressed as

$$\hat{\mathcal{T}} = \sum_{\mathbf{k}} \sum_{mm'} \epsilon^{mm'}_{\mathbf{k}\sigma} c^\dagger_{\mathbf{k}m\sigma} c_{\mathbf{k}m'\sigma}\,, \tag{3.10}$$

where $c^\dagger_{\mathbf{k}m\sigma}$ is the Fourier transform of $c^\dagger_{\mathbf{R}m\sigma}$, and the dispersion $\epsilon^{mm'}_{\mathbf{k}\sigma}$ is given by the transform of the Bloch dispersion $\epsilon_{\mathbf{k}n\sigma}$

$$\epsilon^{mm'}_{\mathbf{k}\sigma} = \sum_{n\in\mathcal{C}} (\bar{P}^\sigma_{mn}(\mathbf{k}))^* \epsilon_{\mathbf{k}n\sigma} \bar{P}^\sigma_{m'n}(\mathbf{k})\,. \tag{3.11}$$

This concludes the discussion on the single-particle part of the multi-band generalization of the Hubbard model, yielding local Wannier functions with well defined local orbital angular momentum quantum numbers.

3.2 Multi-band interactions

The next step is to model the electron-electron Coulomb interaction. Due to the localized character of the d-band Wannier functions of transition metals,

the electron interactions are approximately local, and can be modeled as non-zero only when two electrons occupy the same lattice site $\mathbf{R}$. In terms of the general two-particle interaction vertex defined in Eq. (1.9) this amounts to the approximation

$$V_{\alpha\beta\gamma\delta} = (\mathbf{R}_i a, \mathbf{R}_j b | r_{12}^{-1} | \mathbf{R}_k c, \mathbf{R}_l d) \approx \delta_{ijkl} (ab | r_{12}^{-1} | cd) \,, \tag{3.12}$$

where each label a encompass all local quantum numbers $a = (m, \sigma)$. Insertion of the vertex $V_{\alpha\beta\gamma\delta}$ into Eq. (1.7) gives the second-quantized form of the most general site-local electron-electron interaction as

$$\hat{H}_{int} = \sum_{\mathbf{R}} \sum_{abcd} (ab | r_{12}^{-1} | cd) \, c^\dagger_{\mathbf{R}a} c^\dagger_{\mathbf{R}b} c_{\mathbf{R}d} c_{\mathbf{R}c} \,. \tag{3.13}$$

One fruitful interpretation of $\hat{H}_{int}$ is as a two-particle scattering process. Each term in Eq. (3.13) scatters the two-particle state $c^\dagger_{\mathbf{R}c} c^\dagger_{\mathbf{R}d} |0\rangle$ with amplitude $V_{abcd} = (ab|r_{12}^{-1}|cd)$ into the two-particle state $c^\dagger_{\mathbf{R}a} c^\dagger_{\mathbf{R}b} |0\rangle$. From here on we will drop the site label $\mathbf{R}$ on fermion operators when dealing with site-local forms of interactions.

In contrast to the single-band Hubbard model with its single interaction parameter U we now have to deal with a potentially large number of independent scattering amplitudes V_{abcd}. However not all values V_{abcd} are independent, due to symmetry, as the local interaction must be invariant with respect to the operations of the lattice point group. We are not going to develop the theory for point group symmetric interactions here, although it would be an interesting route. Instead, we will approximate the local interaction as being rotationally invariant, a good approximation for many transition metal systems [37]. This assumption also enables us to reuse much formalism and results from atomic many-body theory [38–40].

3.3 Rotational invariance

An operator $\hat{\mathcal{O}}$ that, under any rotation in spin or angular momentum space, remains unaltered, is defined as spin and angular momentum rotationally invariant.[2] Both rotation groups are continuous and form two Lie algebras with generators $\hat{L}_i$ and $\hat{S}_i$ $(i = x, y, z)$. Given $\hat{\mathcal{O}}$, the condition for invariance is commutation with the generators of rotations [41]

$$[\hat{L}_i, \hat{\mathcal{O}}] = 0 \,, \quad [\hat{S}_i, \hat{\mathcal{O}}] = 0 \,. \tag{3.14}$$

[2]Note that this implies imposing two separate symmetries wrt. $\mathbf{S}$ and $\mathbf{L}$ separately, as opposed to invariance only in total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

Thus it is straightforward to test rotational invariance. The commutation relations also show that for an operator to be invariant it must transform as if having zero total spin and angular momentum $S = L = 0$. We will now see how such operators can be constructed.

As an example, take a double tensor operator $T^{(LS)}_{M\Sigma}$, transforming like a spherical harmonic function with total L and z-component M in angular momentum and as a spin function with total S and z-component Σ in spin [39]. Rotationally invariant operators can then be constructed by the scalar product [38]

$$T^{(LS)} \cdot T^{(LS)} \equiv (-1)^{L+S}\sqrt{(2L+1)(2S+1)}\,\{T^{(LS)}T^{(LS)}\}^{(00)}_{00}\,, \qquad (3.15)$$

defined as the vector coupling $\{...\}^{(LS)}$ with zero total spin and angular momentum $L = S = 0$. It is possible to construct invariant two-particle scattering operators in exactly the same way [42].

Consider the fermion creation operator $c^\dagger_{m\sigma}$ for a fixed single particle angular momentum l and spin $s = 1/2$. Studying the commutation relations with the generators $\hat{L}_i$ and $\hat{S}_i$ then gives, by definition [Eqs. (3.8) and (3.9)]

$$[\hat{L}_z, c^\dagger_{m\sigma}] = m c^\dagger_{m\sigma}\,, \qquad (3.16)$$

$$[\hat{L}_\pm, c^\dagger_{m\sigma}] = \sqrt{l(l+1) - m(m\pm 1)}\, c^\dagger_{m\pm 1,\sigma}\,, \qquad (3.17)$$

$$[\hat{S}_z, c^\dagger_{m\sigma}] = \sigma c^\dagger_{m\sigma}\,, \qquad (3.18)$$

$$[\hat{S}_\pm, c^\dagger_{m\sigma}] = \sqrt{s(s+1) - \sigma(\sigma\pm 1)}\, c^\dagger_{m,\sigma\pm 1}\,. \qquad (3.19)$$

This implies that $c^\dagger_{m\sigma}$ is a double tensor operator of rank s in spin and l in orbital momentum. The same relations holds for the annihilation operator $\tilde{c}_{m\sigma}$ constructed as

$$\tilde{c}_{m\sigma} = (-1)^{s+l-\sigma-m} c_{-m,-\sigma}\,. \qquad (3.20)$$

Now equipped with the single-particle double tensor operators

$$(\mathbf{c}^\dagger)^{(ls)}_{m\sigma} = c^\dagger_{m\sigma}\,, \quad (\mathbf{c})^{(ls)}_{m\sigma} = \tilde{c}_{m\sigma}\,, \qquad (3.21)$$

we can directly construct vector coupled two-particle double tensor operators with fixed total L and S

$$(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(LS)}_{M\Sigma} = \{(\mathbf{c}^\dagger)^{(ls)}(\mathbf{c}^\dagger)^{(ls)}\}^{(LS)}\,, \qquad (3.22)$$

$$(\mathbf{c}\,\mathbf{c})^{(LS)}_{M\Sigma} = \{(\mathbf{c})^{(ls)}(\mathbf{c})^{(ls)}\}^{(LS)}\,. \qquad (3.23)$$

The operators $(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(LS)}_{M\Sigma}(\mathbf{c}\,\mathbf{c})^{(LS)}_{M'\Sigma'}$ form an alternative complete basis for the general two particle operator in Eq. (3.13). However the gain of the double tensor form is that all rotationally invariant combinations are, as in Eq. (3.15), given by the scalar product

$$(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(LS)} \cdot (\mathbf{c}\,\mathbf{c})^{(LS)} \,. \tag{3.24}$$

Thus the most general rotational invariant two-particle operator $\hat{\mathcal{O}}$ can be written as [42]

$$\hat{\mathcal{O}} = -\frac{1}{2}\sum_{LS}(-1)^{L+S}(llLS||\hat{\mathcal{O}}||llLS)\,(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(LS)} \cdot (\mathbf{c}\,\mathbf{c})^{(LS)} \,, \tag{3.25}$$

where the prefactors $(llLS||\hat{\mathcal{O}}||llLS)$ are called the reduced matrix elements of $\hat{\mathcal{O}}$ and completely defines the operator. Thus, imposing rotational invariance reduces the number of parameters defining the scattering amplitudes V_{abcd} of the local interaction $\hat{H}_{loc}$ to the reduced matrix elements $(llLS||\hat{H}_{loc}||llLS)$.

As the electron-electron Coulomb interaction is spin *independent*, its reduced matrix elements only depends on L. For this case we can construct the most general rotational invariant two-particle operators as

$$\hat{\mathcal{O}} = \sum_{L}(llL||\hat{\mathcal{O}}||llL)\hat{P}_L \,, \tag{3.26}$$

where operators $\hat{P}_L$ are constructed from the two-particle scalar products in Eq. (3.25) summed over S

$$\begin{aligned}\hat{P}_L &= -\frac{1}{2}\sum_{S}(-1)^{L+S}(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(LS)} \cdot (\mathbf{c}\,\mathbf{c})^{(LS)} \\ &= \frac{1}{2}\sum_{\sigma\sigma'} c^\dagger_{m_a\sigma}c^\dagger_{m_b\sigma'}c_{m_d\sigma'}c_{m_c\sigma}\left[\sum_{M}\langle l,m_a;l,m_b|LM\rangle\langle LM|l,m_c;l,m_d\rangle\right] ,\end{aligned} \tag{3.27}$$

where the factors $\langle lm_a;l,m_b|LM\rangle$ are Clebsch-Gordan coefficients. For a spin independent rotationally invariant quartic operator $\hat{\mathcal{O}}$ with known single-particle matrix elements $(ab|\mathcal{O}|cd) = \delta_{\sigma_a\sigma_c}\delta_{\sigma_b\sigma_d}(m_am_b|\mathcal{O}|m_cm_d)$ the corresponding reduced matrix elements can be determined by computing the two particle expectation value $\langle LM|\hat{\mathcal{O}}|LM\rangle$. Choosing $\langle L0|\hat{\mathcal{O}}|L0\rangle$ gives

the particularly simple expression

$$(llL||\mathcal{O}||llL) = \sum_{mm'} \langle L0|l,m;l,-m\rangle (m,-m|\mathcal{O}|m',-m')\langle l,m';l,-m'|L0\rangle \,. \quad (3.28)$$

The reduced set of operators $\hat{P}_L$ and coefficients $(llL||\mathcal{O}||llL)$ has proven to be very useful for comparing different interactions, and will later be applied to the full Coulomb interaction and the simplified Kanamori interaction.

3.4 Coulomb interaction

We now turn to the local rotationally invariant electron-electron Coulomb interaction. When expressed in terms of local angular momentum states we will refer to it as the Slater-Condon interaction [43]. The two-particle matrix elements of Eq. (1.9) has the form

$$(ab|r_{12}^{-1}|cd) = \delta_{\sigma_a\sigma_c}\delta_{\sigma_b\sigma_d} \times \iint d\mathbf{r}_1 d\mathbf{r}_2 \, \bar{\psi}_{m_a}(\mathbf{r}_1)\bar{\psi}_{m_b}(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\psi_{m_c}(\mathbf{r}_1)\psi_{m_d}(\mathbf{r}_2)\,, \quad (3.29)$$

in the Wannier basis $\psi_{m\sigma}(\mathbf{r})$ with angular momentum l, factorizable in radial $R_l(r)$, angular $Y_m^l(\Omega)$, and spin parts χ_σ

$$\psi_{m\sigma}(\mathbf{r}) = R_l(r)Y_m^l(\Omega)\chi_\sigma \,. \quad (3.30)$$

To compute the matrix elements in Eq. (3.29) the Coulomb pair potential can be written in terms of the multipole expansion [43]

$$\frac{1}{r_{12}} = \sum_{k=0}^{\infty} \frac{r_<^k}{r_>^{k+1}} \frac{4\pi}{2k+1} \sum_{q=-k}^{k} \bar{Y}_q^k(\Omega_1)Y_q^k(\Omega_2)\,, \quad (3.31)$$

where $r_> = \max(r_1,r_2)$, $r_< = \min(r_1,r_2)$, and Ω_i denotes solid angles. In terms of tensor scalar products the expansion takes the form [39]

$$\frac{1}{r_{12}} = \sum_k \frac{r_<^k}{r_>^{k+1}} \left(\mathbf{C}^{(k)}(1)\cdot\mathbf{C}^{(k)}(2)\right)\,, \quad (3.32)$$

using the spherical harmonic tensor $C_q^{(k)} = \sqrt{4\pi/(2k+1)}Y_q^k$. Inserting the multipole expansion in Eq. (3.29) now gives

$$(ab|r_{12}^{-1}|cd) = \delta(\sigma_a,\sigma_c)\delta(\sigma_b,\sigma_d)\sum_k F^{(k)}(ll,ll) \sum_q (-1)^q \langle lm_a|C_{-q}^k|lm_c\rangle\langle lm_b|C_q^k|lm_d\rangle\,, \quad (3.33)$$

where $F^{(k)}$ is the radial Slater-integral

$$F^{(k)}(ll,ll) = \iint dr_1 dr_2\, \bar{R}_l(r_1)\bar{R}_l(r_2)\frac{r_<^k}{r_>^{k+1}}R_l(r_1)R_l(r_2)\,,$$

and the matrix elements of $C_q^{(k)}$ have the form

$$\langle lm|C_q^k|l'm'\rangle = (-1)^m\sqrt{(2l+1)(2l'+1)}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & k & l' \\ -m & q & m' \end{pmatrix}, \quad (3.34)$$

where the last two factors are Wigner 3j symbols. This now completely determines the matrix elements $(ab|r_{12}^{-1}|cd)$ but still contains the full angular dependence. The corresponding reduced matrix elements can be derived using the rules of higher order vector couplings and results in [38, 40]

$$(llL||r_{12}^{-1}||llL) = (-1)^L\sum_k \begin{Bmatrix} l & l & L \\ l & l & k \end{Bmatrix}(l||C^{(k)}||l)^2F^{(k)}\,, \quad (3.35)$$

where $\{\dots\}$ is a Wigner 6j symbol and $(l||C^{(k)}||l)$ is the reduced matrix element of the spherical harmonic tensor

$$(l||C^{(k)}||l') = (-1)^l\sqrt{(2l+1)(2l'+1)}\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.36)$$

From the properties of the Wigner 3j symbol in Eq. (3.36) all odd factors of k become zero and the triangle identities for the Wigner 6j symbol in Eq. (3.35) limits L and k to $0 \leq L, k \leq 2l$. Thus for d-orbitals with $l = 2$ the only radial integrals $F^{(k)}$ that survives are $F^{(0)}$, $F^{(2)}$ and $F^{(4)}$, which we will treat as parameters of the model. For the d-band it is also convenient to introduce the rescaled parameters

$$F^{(0)} = F_0\,, \quad F^{(2)} = 49F_2\,, \quad F^{(4)} = 441F_4\,, \quad (3.37)$$

when working with multiplet energies as we will see in the coming sections.

Thus we can express the local Coulomb interaction in the Slater-Condon form $\hat{H}_{SC}$ as a sum of operators $\hat{H}_k$ multiplying the parameters $F^{(k)}$

$$\hat{H}_{SC} = \sum_k F^{(k)} \hat{H}_k \,, \tag{3.38}$$

where $\hat{H}_k$ is given by

$$\hat{H}_k = \sum_L (llL||\hat{H}_k||llL)\hat{P}_L \,, \tag{3.39}$$

with reduced matrix elements

$$(llL||\hat{H}_k||llL) = (-1)^L \left\{ \begin{array}{ccc} l & l & L \\ l & l & k \end{array} \right\} (l||C^{(k)}||l)^2 \,, \tag{3.40}$$

and $\hat{P}_L$ as defined in Eq. (3.27).

To conclude we have derived the Slater-Condon interaction as the resulting rotationally invariant form for the Coulomb interaction. For d-electrons with single-particle angular momentum $l = 2$, we obtain a local interaction Hamiltonian $\hat{H}_{SC}$ parametrized by three parameters $F^{(0)}$, $F^{(2)}$ and $F^{(4)}$.

3.5 Atomic limit

In the atomic limit the kinetic single-particle contribution [Eq. (3.10)] to the Hamiltonian disappears, and we are left with an ensemble of isolated atoms. A description of the system can then be obtained from the eigenstates of the local Slater-Condon Coulomb interaction in Eq. (3.38). The atomic multiplet eigenstates can in principle be computed by a diagonalization in the local Fock-space. However the interpretation of the resulting eigenstates is simplified enormously by using a well suited many-body basis that incorporates all the symmetry properties of the interaction.

As the local Hamiltonian $\hat{H}_{SC}$ is both spin and angular momentum rotationally invariant, it commutes with the corresponding Casimir operators $\hat{S}^2$ and $\hat{L}^2$ and one of the respective spin and angular momentum components, e.g., $\hat{S}_z$ and $\hat{L}_z$. Furthermore, the interaction conserves particle number and therefore commutes with the total number operator $\hat{N}$. Thus, a basis having good quantum numbers N, S, Σ, L, and M automatically block-diagonalizes $\hat{H}_{SC}$, where Σ and M are the z-components of total spin and angular momentum respectively.

Assuming that we are dealing with d-electrons with a single-particle angular momentum $l = 2$, the aforementioned set of quantum numbers are not sufficient to fully enumerate all many-body states. However, in the special case of d-electrons it is possible to use the seniority quantum number ν to separate the remaining states. The seniority quantum number was first introduced by Racah [44] and counts the number of unpaired electrons in the state. For example, a state with $N = 3$, $L = 2$ and $S = 1/2$ can be constructed from a single unpaired electron with $l = 2$ and $s = 1/2$ and an electron pair with $L = 0$ and $S = 0$ producing a state with seniority $\nu = 1$. In terms of operators, ν is related to the quasi-spin generators [40]

$$\begin{aligned}
\hat{Q}_1^{(1)} &= -\frac{i}{2}\sqrt{2l+1}(\mathbf{c}^\dagger\mathbf{c}^\dagger)^{(00)}\,, \\
\hat{Q}_{-1}^{(1)} &= -\frac{i}{2}\sqrt{2l+1}(\mathbf{c}\,\mathbf{c})^{(00)}\,, \\
\hat{Q}_0^{(1)} &= -\frac{i}{2\sqrt{2}}\sqrt{2l+1}\left((\mathbf{c}\,\mathbf{c}^\dagger)^{(00)} + (\mathbf{c}^\dagger\mathbf{c})^{(00)}\right)\,,
\end{aligned} \tag{3.41}$$

and their corresponding Casimir operator

$$\hat{Q}^2 = 2\hat{Q}_1^{(1)}\hat{Q}_{-1}^{(1)} + \hat{Q}_z(\hat{Q}_z - 1)\,, \tag{3.42}$$

where $\hat{Q}_z = -i\hat{Q}_0^{(1)} = (\hat{N} - 2l - 1)/2$. The eigenvalues of $\hat{Q}^2$ are $Q(Q+1)$ where Q depends on the seniority quantum number as

$$Q = \frac{1}{2}(2l + 1 - \nu)\,, \tag{3.43}$$

see Ref. [40] for more details.

Studying $\hat{Q}^2$ we see that the only non-trivial part is generated by the first raising and lowering factor, and combining Eq. (3.41) with the definition of the two-particle scattering operators in Eq. (3.27) this term can be written as

$$2\hat{Q}_1^{(1)}\hat{Q}_{-1}^{(1)} = (2l+1)\hat{P}_{L=0}\,, \tag{3.44}$$

where $\hat{P}_{L=0}$ scatters paired electrons with $L = 0$ and $S = 0$, and counts the number of possible pairings of such two-electron pairs.

Returning to the construction of a many-body basis for the d-shell: by simultaneously diagonalizing the operators $\hat{N}$, $\hat{L}^2$, $\hat{L}_z$, $\hat{S}^2$, $\hat{S}_z$, and $\hat{Q}^2$, we obtain a complete basis on the form $|NLMS\Sigma\nu\rangle$. Casting $\hat{H}_{SC}$ in this basis gives a block-diagonal structure in all but the seniority quantum number.

$$\langle NLMS\Sigma\nu|\hat{H}_{SC}|NLMS\Sigma\nu'\rangle = (H_{SC})_{\nu\nu'}^{(NLS)}\,. \tag{3.45}$$

The resulting matrices are at most 3×3 and can readily be diagonalized algebraically, if the freedom in phase is appropriately treated. The multiplet eigenvalues obtained in this way are listed in Appendix A. The agreement between our calculation in the two and tree particle sectors with Ref. [43] and for all particle sector ground states with Ref. [45], is a good sanity check of our implementation of the Slater-Condon interaction.

3.6 Kanamori interaction

For compounds with octahedrally coordinated transition metal atoms, the local crystal field splits the local d-electron single-particle states into two irreducible sets, the t_{2g} (triply degenerate) and e_g (double degenerate) representations. For strong enough crystal field the low energy physics is isolated to only one of the representations and it is possible to construct a low energy model involving only states from that irreducible representation. The corresponding reduction of the Slater-Condon interaction to just one irreducible representation was first proposed by Kanamori [46], to describe ferromagnetic metals.

Before writing down the interaction we first define a diagonalizing unitary transform of the spherical harmonics that block-diagonalizes the single-particle t_{2g} and e_g irreducible representations. One such basis is the cubic harmonics $\mathcal{Y}^l_\alpha$ that are composed by all real linear combinations of spherical harmonics Y^l_m [43]

$$\mathcal{Y}^l_\alpha(\Omega) = \begin{cases} \frac{1}{\sqrt{2}}(Y^l_{-\alpha} + (-1)^\alpha Y^l_\alpha)\,, & \alpha > 0 \\ Y^l_0\,, & \alpha = 0 \\ \frac{1}{i\sqrt{2}}(Y^l_{-\alpha} - (-1)^\alpha Y^l_\alpha)\,, & \alpha < 0 \end{cases} . \tag{3.46}$$

Collecting the transformation rules into the unitary transform $\mathcal{U}_{\alpha m}$ gives the corresponding cubic harmonic second-quantization operators $c_{\alpha\sigma}$ in terms of the previously defined spherical harmonic fermion operators $c_{m\sigma}$

$$c_{\alpha\sigma} = \sum_m \mathcal{U}_{\alpha m} c_{m\sigma}\,. \tag{3.47}$$

In condensed matter physics the cubic harmonics are often called atomic orbitals and discussed in terms of their Cartesian dependencies, which for the d-electrons are, $x^2 - y^2 \propto \mathcal{Y}^2_2$, $xy \propto \mathcal{Y}^2_1$, $3z^2 - r^2 \propto \mathcal{Y}^2_0$, $yz \propto \mathcal{Y}^2_{-1}$,

and $xy \propto \mathcal{Y}^2_{-2}$. The octahedral symmetry group divides these states in the irreducible representations $xy, xy, yz \in t_{2g}$ and $x^2 - y^2, 3z^2 - r^2 \in e_g$.

The rotationally invariant Coulomb interaction mapped on to one of the irreducible representations now takes the form of the Kanamori interaction Hamiltonian [47]

$$\begin{aligned}\hat{H}_K = U \sum_{\alpha} \hat{n}_{\alpha\uparrow}\hat{n}_{\alpha\downarrow} + \frac{1}{2}\sum_{\alpha\neq\beta}\sum_{\sigma\sigma'}(U' - J\delta_{\sigma\sigma'})\hat{n}_{\alpha\sigma}\hat{n}_{\beta\sigma'} \\ - \sum_{\alpha\neq\beta}\Big(J\underbrace{c^\dagger_{\alpha\uparrow}c_{\alpha\downarrow}c^\dagger_{\beta\downarrow}c_{\beta\uparrow}}_{\text{spin-flip}} + J'\underbrace{c^\dagger_{\beta\uparrow}c^\dagger_{\beta\downarrow}c_{\alpha\uparrow}c_{\alpha\downarrow}}_{\text{pair-hopping}}\Big), \quad (3.48)\end{aligned}$$

where the interaction parameters are given by the matrix elements

$$\begin{aligned} U &= (\alpha\alpha|r_{12}^{-1}|\alpha\alpha)\,, \quad U' = (\alpha\beta|r_{12}^{-1}|\alpha\beta) \\ J &= (\alpha\beta|r_{12}^{-1}|\beta\alpha)\,, \quad J' = (\alpha\alpha|r_{12}^{-1}|\beta\beta)\,, \end{aligned} \quad (3.49)$$

that depend on each other as $U' = U - 2J$, $J' = J$ due to rotational invariance.

The interaction terms of the Kanamori Hamiltonian [Eq. (3.48)] are divided up, from the beginning to the end, in intra-orbital density-density interaction, inter-orbital density-density interaction, spin-flip interaction, and pair-hopping interaction. Among the terms the pair-hopping is interesting as it scatters spin paired two-particle states in the same cubic harmonic state to a pair in another cubic harmonic state. Such scattering could be suspected to not conserve angular momentum, but it actually does. This stems from that the cubic harmonic states themselves carry a net zero angular momentum, as they are linear combinations of spherical harmonic states with m and $-m$ angular momentum z-components.

From the form of the Kanamori interaction in Eq. (3.48) the spin and rotational invariance is not directly evident, but the symmetry becomes apparent when normal ordering the interaction. In fact, all terms can be expressed using normal ordered operators on the form

$$\hat{\mathcal{O}}_{\alpha\beta\gamma\delta} = \frac{1}{2}\sum_{\sigma\sigma'} c^\dagger_{\alpha\sigma}c^\dagger_{\beta\sigma'}c_{\delta\sigma'}c_{\gamma\sigma}\,, \quad (3.50)$$

where the two-particle scattering is explicitly spin independent, and thus spin rotationally invariant.

In terms of the $\hat{\mathcal{O}}$ operators, $\hat{H}_K$ separates into

$$\hat{H}_K = (U - 2J)\sum_{\alpha\beta}\hat{\mathcal{O}}_{\alpha\beta\alpha\beta} + J\sum_{\alpha\beta}\Big(\hat{\mathcal{O}}_{\alpha\beta\beta\alpha} + \hat{\mathcal{O}}_{\alpha\alpha\beta\beta}\Big)\,, \quad (3.51)$$

where the first scattering term is a general orbital preserving scattering, the second is an orbital exchange scattering, and the last is a orbital pair scattering term.

Now the total spin Casimir operator can be expressed as

$$\hat{S}^2 = \sum_{ab} \left[-\frac{1}{2}\hat{\mathcal{O}}_{\alpha\beta\alpha\beta} - \hat{\mathcal{O}}_{\alpha\beta\beta\alpha} \right] + \frac{3}{4}\hat{N}\,, \tag{3.52}$$

and the raising and lowering term of Eq. (3.44), contained in the expression in Eq. (3.42) for the quasi-spin Casimir operator $\hat{Q}^2$, is equal to

$$(2l+1)\hat{P}_{L=0} = \sum_{\alpha\beta} \hat{\mathcal{O}}_{\alpha\alpha\beta\beta}\,. \tag{3.53}$$

Using these relations the Kanamori interaction can be written in terms of spin and quasi-spin as

$$\hat{H}_K = (U-3J)\frac{1}{2}\hat{N}(\hat{N}-1) + J\left(\hat{Q}^2 - \hat{S}^2\right)\,, \tag{3.54}$$

neglecting an irrelevant constant shift, and a $\hat{N}$ dependent chemical potential shift.

Computing the matrix elements in Eq. (3.49) from the Slater-Condon form of the Coulomb interaction in Eq. (3.38) we obtain

$$U_{t_{2g}} = F_0 + 4F_2 + 36F_4\,, \tag{3.55}$$
$$J_{t_{2g}} = 3F_2 + 20F_4\,, \tag{3.56}$$

for the t_{2g} states, and

$$U_{e_g} = F_0 + 4F_2 + 36F_4\,, \tag{3.57}$$
$$J_{e_g} = 4F_2 + 15F_4\,. \tag{3.58}$$

for the e_g states, with F_0, F_2, and F_4 as defined in Eq. (3.37), in agreement with Ref. [9].

The focus of this thesis is the full d-band, so in a sense the specialization of the Kanamori model for the reduced t_{2g} and e_g model is not directly relevant. Our current interest in the Kanamori interaction is motivated by a series of studies [47–53], applying the Kanamori interaction [Eq. (3.48)] to the *full* set of d-states. For this case the Kanamori interaction, having one interaction parameter less than the full Slater-Condon interaction, can not be a complete general description of the Coulomb interaction. It is not even

clear whether the Kanamori interaction applied to the full d-band is contained within the possible configurations of the Slater-Condon interaction, or not.

However from the t_{2g} and e_g mappings we see that the only point where the Kanamori interaction correctly describes the interactions within both irreducible representations is when $J_{t_{2g}} = J_{e_g}$, i.e. when $F^{(4)}/F^{(2)} = 9/5$. But we do not know a priori whether the t_{2g} – e_g interaction in the Slater-Condon interaction at $F^{(4)}/F^{(2)} = 9/5$, is reduced to the Kanamori form. To investigate this we turn to an analysis of the reduced matrix elements of both interactions.

3.7 Reduced matrix elements

The structure of the Kanamori and Slater-Condon interactions applied to the full set of d-orbitals can be analyzed due to the fact that both interactions are rotationally invariant, expressible in the form of Eq. (3.26), and thus completely determined by their respective reduced matrix elements.

To compute the reduced matrix elements of the Kanamori interaction $\hat{H}_K$ the single particle operators have to be transformed, from the cubic harmonics used in Eq. (3.48) to the spherical-harmonic basis using the unitary transform in Eq. (3.47). The two degrees of freedom in the coupling parameters U and J can be separated giving

$$\hat{H}_K = U\hat{H}_U + J\hat{H}_J\,. \tag{3.59}$$

In the spherical harmonic basis, Eq. (3.28) is then used to compute the reduced matrix elements of the two operators $(llL||\hat{H}_U||llL)$ and $(llL||\hat{H}_J||llL)$. For the Slater-Condon interaction $\hat{H}_{SC}$, Eq. (3.40) directly gives the reduced matrix elements of all three operators $\hat{H}_k$ multiplying the coupling parameters $F^{(0)}$, $F^{(2)}$, and $F^{(4)}$. The results are listed in Table 3.1, where the different values of L are denoted by the corresponding two particle multiplets on the form ^{2S+1}L.

The reduced matrix elements in Table 3.1 are vector representations of the two-particle operators in the five dimensional vector-space of two-particle multiplets. For an operator $\hat{\mathcal{O}}$, the corresponding reduced matrix element vector is given by

$$\mathbf{v}(\hat{\mathcal{O}}) = [(ll0||\hat{\mathcal{O}}||ll0), \ldots, (ll(2l)||\hat{\mathcal{O}}||ll(2l))]\,. \tag{3.60}$$

The interactions $\hat{H}_K$ and $\hat{H}_{SC}$ then reduce to two and three dimensional

Table 3.1: Reduced matrix elements $(llL||\hat{\mathcal{O}}||llL)$ for the Hamiltonian components $\hat{H}_U$ and $\hat{H}_J$ in the Kanamori interaction [Eq. (3.59)], and the Slater-Condon interaction components $\hat{H}_1$, $\hat{H}_2$, and $\hat{H}_4$ [Eq. (3.39)], for the d-shell ($l = 2$). Note the scaling of the Slater-Condon operators.

$l = 2$	Kanamori		Slater-Condon			Operators	
^{2S+1}L	$\hat{H}_U$	$\hat{H}_J$	$\hat{H}_0$	$49\hat{H}_2$	$441\hat{H}_4$	$\hat{N}^2/2$	$(\hat{Q}^2 - \hat{S}^2)$
^1S	1	4	1	14	126	1	7
^3P	1	−3	1	7	−84	1	0
^1D	1	−1	1	−3	36	1	2
^3F	1	−3	1	−8	−9	1	0
^1G	1	−1	1	4	1	1	2

linear combinations of reduced matrix element vectors

$$\mathbf{v}(\hat{H}_K) = U\mathbf{v}(\hat{H}_U) + J\mathbf{v}(\hat{H}_J)\,, \tag{3.61}$$

$$\mathbf{v}(\hat{H}_{SC}) = F^{(0)}\mathbf{v}(\hat{H}_0) + F^{(2)}\mathbf{v}(\hat{H}_2) + F^{(4)}\mathbf{v}(\hat{H}_4)\,. \tag{3.62}$$

The question whether the Kanamori interaction can be faithfully described by the full Slater-Condon interaction can now be reformulated as whether the simple vector equations

$$[\mathbf{v}(\hat{H}_0), \mathbf{v}(\hat{H}_2), \mathbf{v}(\hat{H}_4)]\mathbf{x}_U = \mathbf{v}(\hat{H}_U)\,, \tag{3.63}$$

$$[\mathbf{v}(\hat{H}_0), \mathbf{v}(\hat{H}_2), \mathbf{v}(\hat{H}_4)]\mathbf{x}_J = \mathbf{v}(\hat{H}_J)\,, \tag{3.64}$$

are solvable. It turns out that they are, which implies that the Kanamori interaction spans a subspace of the Slater-Condon interaction

$$\mathrm{span}\{\mathbf{v}(\hat{H}_U), \mathbf{v}(\hat{H}_J)\} \subset \mathrm{span}\{\mathbf{v}(\hat{H}_0), \mathbf{v}(\hat{H}_2), \mathbf{v}(\hat{H}_4)\}\,. \tag{3.65}$$

The solution gives that the Slater-Condon interaction applied to the d-shell is identically equal to the Kanamori interaction when

$$F^{(0)} = U - \frac{8}{5}J\,, \quad F^{(2)} = 7J\,, \quad F^{(4)} = \frac{63}{5}J\,, \tag{3.66}$$

where the ratio of $F^{(4)}$ and $F^{(2)}$ is fixed to

$$\frac{F^{(4)}}{F^{(2)}} = \frac{9}{5}\,. \tag{3.67}$$

Thus at this fixed ratio, we have proved that, in addition to our previous observation the interaction within the t_{2g} and e_g are the same for this ratio, and also the $e_g - t_{2g}$ interaction has exactly the Kanamori form of Eq. (3.48).

This observation now enables us to seek a parametrization of the Slater parameters $F^{(k)}$ in terms of the Kanamori parameters U and J plus an extra parameter that that control the departure from the Kanamori interaction case of Eq. (3.66).

3.8 Slater-Condon (U, J) parametrization

The prospect of using U and J to determine $F^{(k)}$ is interesting in part due to the fact that many aspects of the Hubbard U and Hund's rule J are already understood from studies of two and three band models [9], but also for comparison with previous studies of the full d-band using the Kanamori interaction [47–53].

How to introduce a third parameter to regain the full three parameter freedom of the Slater-Condon interaction is in a sense arbitrary, but forced to make a choice we have constructed a parametrization based on preexisting formalism. Let us first introduce the Racah parameters [54]

$$A = F_0 - 49F_4\,, \quad B = F_2 - 5F_4\,, \quad C = 35F_4\,. \tag{3.68}$$

This is a one-to-one alternate parametrization to the Slater parameters $F^{(k)}$. In terms of the Racah parameters the Kanamori limit given in Eq. (3.66) takes the form

$$A = U - 3J\,, \quad B = 0\,, \quad C = J\,. \tag{3.69}$$

Note that B turns out to be zero for the Kanamori interaction. Looking further into the early literature on the Slater-Condon interaction, Tanabe-Sugano [55] introduced the ratio $C/B = \gamma$ and noted that it is nearly constant for all elements, $\gamma \approx 4 - 5$. With this background we suggest the use of $1/\gamma$ as the third parameter, which gives

$$A = U - 3J\,, \quad B = J/\gamma\,, \quad C = J\,. \tag{3.70}$$

Translated back to the Slater parameters $F^{(k)}$, the introduction of the

Tanabe-Sugano ratio γ gives

$$\begin{aligned} F^{(0)} &= U - \frac{8}{5}J\,, \\ F^{(2)} &= 49\left(\frac{1}{\gamma} + \frac{1}{7}\right)J\,, \\ F^{(4)} &= \frac{63}{5}J\,. \end{aligned} \tag{3.71}$$

Tanabe-Sugano's observation that γ is approximately constant directly translates into the $F^{(4)}/F^{(2)}$ ratio

$$\frac{F^{(4)}}{F^{(2)}} = 9\frac{F_4}{F_2} = \frac{9}{5}\frac{1}{7/\gamma + 1}\,. \tag{3.72}$$

The fact that the $F^{(4)}/F^{(2)}$ ratio varies weakly for the transition metal elements has been reported by several authors [37, 55, 56], and the reported range is $F^{(4)}/F^{(2)} \approx 0.65$ – 0.85. The fact that the Kanamori interaction applied to the full d-shell corresponds to the slightly unphysical ratio $F^{(4)}/F^{(2)} = 1.8$, sheds doubt on its applicability as interaction model for this case.

The point in the parameter space of the Slater-Condon interaction, corresponding to $F^{(4)}/F^{(2)} = 9/5$ (or equivalently $1/\gamma = 0$), has previously been noted as special by Judd *et al.* [57], who identified it as a so called Laporte-Platt degenerate point with large accidental degeneracies in the multiplet spectra. However, to the best of our knowledge, the connection to the Kanamori interaction has not previously been shown.

The degree of accidental degeneracy at $1/\gamma = 0$ can directly be inferred from the spin and quasi-spin form of the Kanamori interaction in Eq. (3.54). The interaction only split multiplets with respect to total spin S and quasi-spin Q, or equivalently the seniority quantum number ν. Thus, as opposed to the general Slater-Condon interaction that do not commute with the quasi-spin Casimir operator, at the Laporte-Platt point, seniority also becomes a good quantum number. At this point the interaction becomes degenerate in total angular momentum L, leading to large accidental degeneracies. In Figure 3.1 the multiplet energies for all fillings are plotted as a function of $1/\gamma$, and it is directly evident that multiplets with the same spin S and seniority ν becomes degenerate at $1/\gamma = 0$, independent of their angular momentum quantum numbers L.

The motivation and derivation of the Kanamori-Tanabe-Sugano parametrization of the Slater-Condon interaction is a central theme in Paper V,

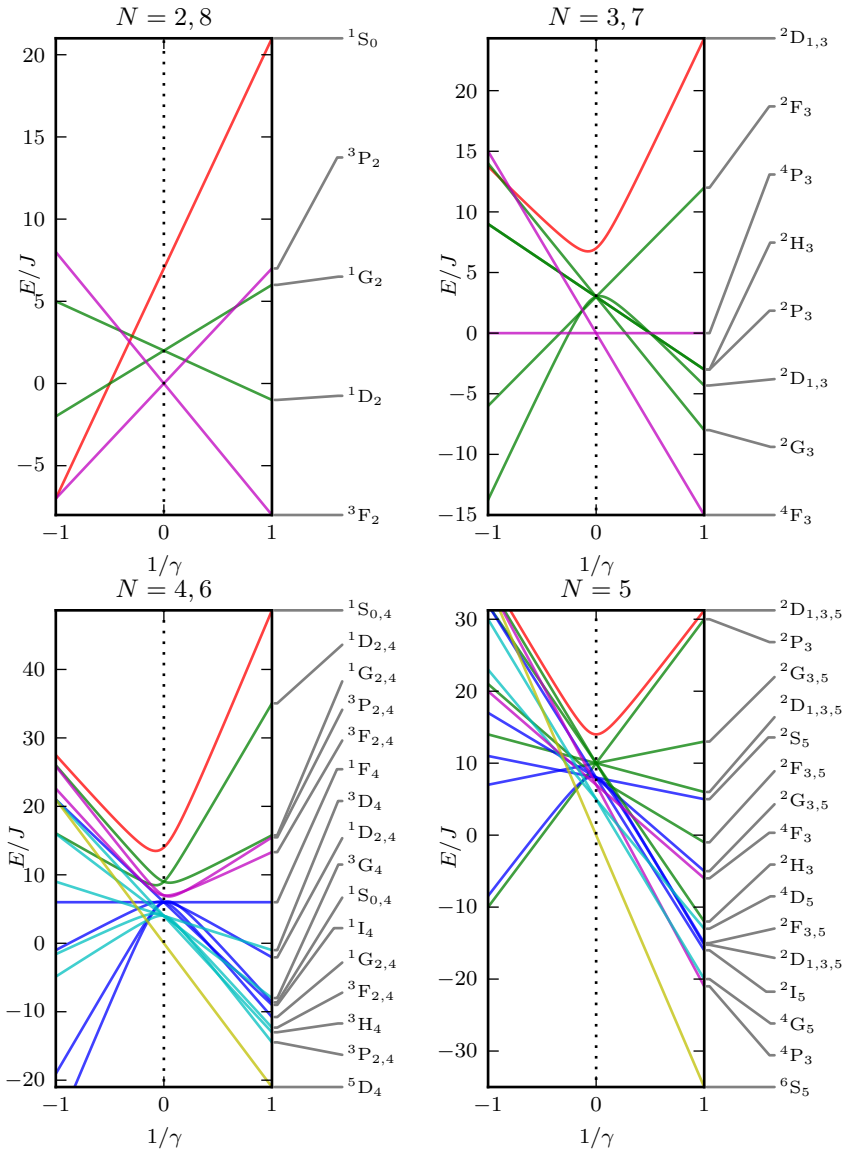


Figure 3.1: Slater-Condon multiplet energies E as a function of $1/\gamma$, for $U = 3$ and $J/U = 1/3$. The multiplet labels indicate the total spin S, angular momentum L, and seniority ν, on the form $^{2S+1}L_\nu$. The multiplets with same spin S and seniority ν at $1/\gamma = 0$ share color.

where it is used to investigate how the paramagnetic phase boundaries of the d-band Hubbard model differ between the Kanamori limit $1/\gamma = 0$ and the more physical range $1/\gamma = 1/4$, corresponding to $F^{(4)}/F^{(2)} \approx 0.65$. For some integer fillings, the Kanamori model has an accidental degeneracy of two multiplets even for the ground state. The effect of this ground state degeneracy is also investigated in Paper V.

In Paper VI, the focus is instead put on the generated effective Hubbard repulsion, and the ability of the Slater-Condon interaction to stabilize valence-skipping charge disproportionate ground states for finite values of $1/\gamma$.

3.9 Parameters U and J in the literature

In the literature there is wide a spread use of U and J for other purposes than as the coupling parameters entering in the Kanamori interaction [Eq. (3.48)]. One such example is the LDA+U formalism that models the local interaction in terms of density-density interactions only. For this type of Ising like interaction the orbital averaged Hubbard $\bar{U}_{orb}$ and Hund's rule $\bar{J}_{orb}$ are used as effective interaction parameters [58]. In terms of the Slater-Condon interaction the orbital averages gives

$$\bar{U}_{orb} = F^{(0)}\,, \quad \bar{J}_{orb} = \frac{1}{14}(F^{(2)} + F^{(4)})\,. \tag{3.73}$$

These averages then enters as coupling parameters in the density-density interaction

$$\hat{H}_{\text{Ising}} = \bar{U}_{orb}\sum_{m}\hat{n}_{m\uparrow}\hat{n}_{m\downarrow} + \frac{1}{2}\sum_{m\neq m'}\sum_{\sigma\sigma'}(\bar{U}_{orb} - \delta_{\sigma\sigma'}\bar{J}_{orb})\hat{n}_{m\sigma}\hat{n}_{m'\sigma'}\,. \tag{3.74}$$

Note however that this is an approximation of the Slater-Condon interaction, while the Kanamori interaction of Eq. (3.48) together with Eq. (3.66) is an exact form of the Slater-Condon interaction when $F^{(4)}/F^{(2)} = 9/5$. To make this point clear we compute the orbital averages of the Slater-Condon interaction at the Kanamori point in terms of the Kanamori Hubbard U and Hund's rule J

$$\bar{U}_{orb} = U - \frac{8}{5}J\,, \quad \bar{J}_{orb} = \frac{7}{5}J\,, \tag{3.75}$$

which clearly shows that the orbital averages and the Kanamori parameters are not the same. It is not uncommon that $\bar{U}_{orb}$ and $\bar{J}_{orb}$ are simply denoted by U and J in the literature [37], in conflict with Kanamoris original

definition [46]. So interpreting general statements on magnitudes of U and J requires explicit knowledge of the applied interaction model. In addition, the relevant values are also dependent on the choice of Wannier function construction, making effective interaction parameters non-transferable even between studies of the same compound [37].

Another occurring use of U and J are to denote the multiplet averaged Hubbard and Hund's rule interactions [59]. This type of averages gives yet another set of expressions in terms of $F^{(k)}$

$$\bar{U}_{mult} = F^{(0)} - \frac{2}{63}(F^{(2)} + F^{(4)})\,, \quad \bar{J}_{mult} = \frac{1}{14}(F^{(2)} + F^{(4)})\,, \tag{3.76}$$

that translate to the relations

$$\bar{U}_{mult} = U - \frac{20}{9}J\,, \quad \bar{J}_{mult} = \frac{7}{5}J\,. \tag{3.77}$$

for the Kanamori U and J at $F^{(4)}/F^{(2)} = 9/5$.

Part II

Methods

Chapter 4

Dynamical Mean Field Theory

The manifold of approximation methods for many-body systems are by tradition composed mainly of perturbation expansions in some small parameter of the system Hamiltonian. This approach generates approximations that capture the physics in only a part of the accessible parameter space and are therefore ill suited for systems with competing energy-scales. Dynamical Mean Field Theory (DMFT) [17, 60] takes a slightly different approach. Instead of considering the Hamiltonian parameters it is possible to make an expansion in the lattice coordination number z. Keeping only the zeroth order term in such and expansion gives a simplified but still non-trivial theory that is exact in the limit $z \to \infty$, for any parameter values.

In this chapter we give an introduction to the limit of infinite coordination number and its implications. The DMFT equations are derived using the cavity construction [17] in Section 4.1, and the set of self-consistent equations are determined in Section 4.2. The self-consistency is readily formulated in terms of a fix-point problem, therefore some basic methods for solving this class of problems are reviewed in Section 4.3.

For the Hubbard model in Eq. (2.4) the local two body electron-electron interaction makes exact evaluation of the self-energy a hard problem. Although weak coupling expansions are possible, going to higher order is encumbered with an increasing number of Brillouin zone integrations. An important simplification of this problem is obtained in the limit of infinite coordination number $z \to \infty$, when applying the appropriate scaling of the hopping parameter $t \to \frac{t}{\sqrt{z}}$. For models with only local interactions like the

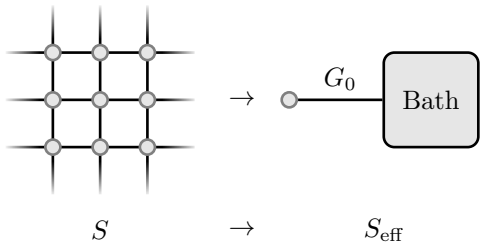


Figure 4.1: The lattice problem and associated auxiliary impurity problem.

Hubbard model, this scaling makes momentum conservation irrelevant in the interaction vertices of all self-energy diagrams and makes the self-energy Σ momentum independent [29, 61, 62]

$$\Sigma(\mathbf{k}, i\omega_n) \underset{z\to\infty}{\longrightarrow} \Sigma(i\omega_n)\,. \tag{4.1}$$

The corresponding interacting lattice Green's function

$$G_L(\mathbf{k}, i\omega_n) = \frac{1}{i\omega_n - \epsilon_{\mathbf{k}} + \mu - \Sigma(i\omega_n)}\,, \tag{4.2}$$

then only depends on the momentum $\mathbf{k}$ through the non-interacting dispersion $\epsilon_{\mathbf{k}}$. Due to this the local interacting lattice Green's function $G_L(i\omega_n)$ can be calculated from the self-energy $\Sigma(i\omega_n)$ and the non-interacting density of states $\rho^{(0)}(\epsilon)$ as

$$\begin{aligned} G_L(i\omega_n) &= G_L(\mathbf{R} = \mathbf{0}, i\omega_n) \\ &= \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{0}} G_L(\mathbf{k}, i\omega_n) = \int d\epsilon\, \frac{\rho^{(0)}(\epsilon)}{i\omega_n - \epsilon + \mu - \Sigma(i\omega_n)}\,. \end{aligned} \tag{4.3}$$

The locality of the self-energy is the main starting point in the development of DMFT. The basic idea is to introduce an auxiliary impurity problem, with the same local interaction as the lattice, connected to a non-interacting bath. Requiring that the impurity self-energy is equal to the lattice self energy gives the self-consistency condition

$$G_0^{-1} = G_L^{-1} + \Sigma\,, \tag{4.4}$$

where $G_0(i\omega_n)$ is the bath Green's function. This procedure amounts to map the action S of the lattice problem to an effective impurity action S_{eff} as depicted in Fig. 4.1.

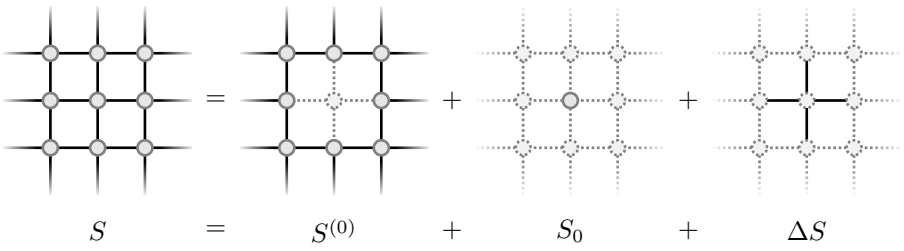


Figure 4.2: The cavity construction.

One of the possible routes to derive the impurity problem and the self-consistency condition is through the cavity construction [17], where the action of the lattice problem S is separated in three parts. By removing one site and its connection to the rest of the lattice, a lattice with a cavity $S^{(0)}$ is formed. The removed parts are the local terms on the cavity site S_0 and its connection to the rest of the lattice ΔS. All in all the separation can be expressed as

$$S = S^{(0)} + S_0 + \Delta S\,, \tag{4.5}$$

see Fig. 4.2 for a schematic picture.

Integrating out all electronic degrees of freedom except the one at the cavity site now gives an effective description of the auxiliary impurity model and the self-consistency Eq. (4.4) for the bath Green's function G_0. In the next section we go through the details of this construction using the coherent state path integral formalism [14, 23].

4.1 Cavity construction

In the coherent-state path-integral formalism the partition function $\mathcal{Z}$ of Eq. (2.5) can be written as

$$\mathcal{Z} = \int \mathcal{D}(\psi^\dagger,\psi) e^{-\int_0^\beta d\tau \left(\sum_{\mathbf{R}\sigma} \psi^\dagger_{\mathbf{R}\sigma}(\tau)(\partial_\tau-\mu)\psi_{\mathbf{R}\sigma}(\tau)+\hat{H}[\psi^\dagger(\tau),\psi(\tau)]\right)}\,, \tag{4.6}$$

where $\psi^\dagger$ and ψ are the Grassman fields of the second quantization creation and annihilation operators, $c^\dagger$ and c, and $\mathcal{D}(\cdot)$ is the standard integration measure of Ref. [14]. In the following we shall make no notational distinction between operators and Grassman fields and use $c^\dagger$ and c to denote both. The

action $S[c^\dagger, c]$ of the system is identified as the expression in the exponent of Eq. (4.6)

$$S[c^\dagger(\tau), c(\tau)] = \int_0^\beta d\tau \left(\sum_{\mathbf{R}\sigma} c^\dagger_{\mathbf{R}\sigma}(\tau)(\partial_\tau - \mu) c_{\mathbf{R}\sigma}(\tau) + \hat{H}[c^\dagger(\tau), c(\tau)] \right) . \tag{4.7}$$

For the Hubbard model given by Eq. (2.4) this becomes

$$S[c^\dagger(\tau), c(\tau)] = \int_0^\beta d\tau \left(\sum_{\mathbf{R}\sigma} c^\dagger_{\mathbf{R}\sigma}(\tau)(\partial_\tau - \mu) c_{\mathbf{R}\sigma}(\tau) + \sum_{\mathbf{R}\mathbf{R}'} \sum_\sigma t_{\mathbf{R}\mathbf{R}'} c^\dagger_{\mathbf{R}\sigma}(\tau) c_{\mathbf{R}'\sigma}(\tau) + U \sum_{\mathbf{R}} \hat{n}_{\mathbf{R}\uparrow}(\tau) \hat{n}_{\mathbf{R}\downarrow}(\tau) \right) , \tag{4.8}$$

where the general hopping-matrix $t_{\mathbf{R}\mathbf{R}'}$ later will be replaced with the nearest-neighbor hopping $t_{\mathbf{R}\mathbf{R}'} = -\delta_{\langle \mathbf{R}\mathbf{R}' \rangle} t$. Now we perform the cavity construction for the site $\mathbf{R} = \mathbf{0}$, and partition the lattice action S in the local part S_0 at $\mathbf{R} = \mathbf{0}$, the action ΔS for all connections between this and all other sites $\mathbf{R} \neq \mathbf{0}$, and the action $S^{(0)}$ for all other sites $\mathbf{R} \neq \mathbf{0}$ as in Eq. (4.5). In terms of the partition function $\mathcal{Z}$ this amounts to

$$\mathcal{Z} = \int \mathcal{D}(c^\dagger, c) e^{-S[c^\dagger, c]} = \int \mathcal{D}(c^\dagger, c) e^{-(S^{(0)} + \Delta S + S_0)} , \tag{4.9}$$

where the actions S_0, ΔS, and $S^{(0)}$ are given by

$$S_0 = \int_0^\beta d\tau \left(\sum_\sigma c^\dagger_{0\sigma}(\tau)(\partial_\tau - \mu) c_{0\sigma}(\tau) + U \hat{n}_{0\uparrow}(\tau) \hat{n}_{0\downarrow}(\tau) \right) , \tag{4.10}$$

$$\Delta S = \int_0^\beta d\tau \sum_{\mathbf{R}\sigma} t_{\mathbf{R}0} \left(c^\dagger_{\mathbf{R}\sigma}(\tau) c_{0\sigma}(\tau) + c^\dagger_{0\sigma}(\tau) c_{\mathbf{R}\sigma}(\tau) \right) , \tag{4.11}$$

$$S^{(0)} = \int_0^\beta d\tau \left(\sum_{\mathbf{R}\neq 0,\sigma} c^\dagger_{\mathbf{R}\sigma}(\tau)(\partial_\tau - \mu) c_{\mathbf{R}\sigma}(\tau) + \sum_{\mathbf{R},\mathbf{R}'\neq 0} \sum_\sigma t_{\mathbf{R}\mathbf{R}'} c^\dagger_{\mathbf{R}\sigma}(\tau) c_{\mathbf{R}'\sigma}(\tau) + U \sum_{\mathbf{R}\neq 0} \hat{n}_{\mathbf{R}\uparrow}(\tau) \hat{n}_{\mathbf{R}\downarrow}(\tau) \right) . \tag{4.12}$$

As the first step in integrating out all sites $\mathbf{R} \neq 0$, we separate out the integration of the cavity site at $\mathbf{R} = 0$

$$\begin{aligned}\mathcal{Z} &= \int \mathcal{D}(c_0^\dagger, c_0) e^{-S_0} \int \prod_{\mathbf{R}\neq 0} \mathcal{D}(c_{\mathbf{R}}^\dagger, c_{\mathbf{R}}) e^{-(S^{(0)}+\Delta S)} \\ &= \int \mathcal{D}(c_0^\dagger, c_0) e^{-S_0} \mathcal{Z}^{(0)} \langle e^{-\Delta S} \rangle_{S^{(0)}} \,. \end{aligned} \tag{4.13}$$

Our aim is to obtain an effective local description of the lattice in terms of an auxiliary impurity problem. For this reason we introduce the effective action S_{eff} of the impurity as

$$\mathcal{Z} = \int \mathcal{D}(c_0^\dagger, c_0) e^{-S_{\text{eff}}} = \int \mathcal{D}(c_0^\dagger, c_0) e^{-S_0} \mathcal{Z}^{(0)} \langle e^{-\Delta S} \rangle_{S^{(0)}} \,. \tag{4.14}$$

Collecting the exponent of the right hand side gives

$$S_{\text{eff}} = S_0 + \ln \langle e^{-\Delta S} \rangle_{S^{(0)}} + \ln \mathcal{Z}^{(0)} \,, \tag{4.15}$$

where S_0 is purely local and the connection to the lattice is described by the logarithm of the expectation value of the hybridization ΔS. Investigating this term in detail reveals that it is possible to rewrite it as an expectation value of a source term with respect to the lattice (with cavity) action $S^{(0)}$

$$\begin{aligned}\ln \langle e^{-\Delta S} \rangle_{S^{(0)}} &= \ln \left[\frac{1}{\mathcal{Z}^{(0)}} \int \prod_{\mathbf{R}\neq 0} \mathcal{D}(c_{\mathbf{R}}^\dagger, c_{\mathbf{R}}) e^{-S^{(0)}} e^{-\Delta S} \right] \\ &= \ln \left\langle e^{-\int_0^\beta d\tau \, \sum_{\mathbf{R}\sigma} t_{\mathbf{R}0} \left(c_{\mathbf{R}\sigma}^\dagger c_{0\sigma} + c_{0\sigma}^\dagger c_{\mathbf{R}\sigma} \right)} \right\rangle_{S^{(0)}} \\ &= \ln \left\langle e^{-\int_0^\beta d\tau \, \sum_{\mathbf{R}\sigma} \left(c_{\mathbf{R}\sigma}^\dagger \eta_{\mathbf{R}\sigma} + \eta_{\mathbf{R}\sigma}^\dagger c_{\mathbf{R}\sigma} \right)} \right\rangle_{S^{(0)}} \,, \end{aligned} \tag{4.16}$$

introducing the local source fields $\eta_{\mathbf{R}\sigma} = t_{\mathbf{R}0} c_{0\sigma}$. In this form Eq. (4.16) takes the exact form of the generating functional $\mathcal{W}^{(0)}(\eta^\dagger, \eta)$ for the connected Green's functions $G_n^{(0)}$ of the action $S^{(0)}$ (to any order n), for a detailed exposé see Ref. [14]

$$\begin{aligned}&\mathcal{W}^{(0)}(\eta^\dagger, \eta) = \ln \left\langle e^{-\int_0^\beta d\tau \, \sum_{\mathbf{R}\sigma} \left(c_{\mathbf{R}\sigma}^\dagger \eta_{\mathbf{R}\sigma} + \eta_{\mathbf{R}\sigma}^\dagger c_{\mathbf{R}\sigma} \right)} \right\rangle_{S^{(0)}} \,, \\ &G^{(0)}_{n;\mathbf{R}_1 \ldots \mathbf{R}_n, \mathbf{R}'_1 \ldots \mathbf{R}'_n}(\tau_1 \ldots \tau_n, \tau'_1 \ldots \tau'_n) = \\ &\qquad (-1)^n \frac{\partial^{2n} \mathcal{W}^{(0)}(\eta^\dagger, \eta)}{\partial \eta^\dagger_{\mathbf{R}_1}(\tau_1) \ldots \partial \eta^\dagger_{\mathbf{R}_n}(\tau_n) \partial \eta_{\mathbf{R}'_1}(\tau'_1) \ldots \partial \eta_{\mathbf{R}'_n}(\tau'_n)} \bigg|_{\eta^\dagger, \eta = 0} \,. \end{aligned} \tag{4.17}$$

With the derivatives of $\mathcal{W}^{(0)}$ known, it is straight forward to write down its perturbation expansion as

$$\mathcal{W}^{(0)}(\eta^\dagger,\eta) = \sum_{n=1} \frac{(-1)^n}{n!} \int_0^\beta d\tau_1 ... d\tau_n d\tau'_1 ... d\tau'_n \times \sum_{i_1 ... i_n} \sum_{j_1 ... j_n} \eta^\dagger_{\mathbf{R}_1}(\tau_1) ... \eta^\dagger_{\mathbf{R}_n}(\tau_n) \eta_{\mathbf{R}'_1}(\tau'_1) ... \eta_{\mathbf{R}'_n}(\tau'_n) \times G^{(0)}_{n;\mathbf{R}_1 ... \mathbf{R}_n, \mathbf{R}'_1 ... \mathbf{R}'_n}(\tau_1 ... \tau_n, \tau'_1 ... \tau'_n) \,. \quad (4.18)$$

Now we are ready to take the limit of infinite connectivity z in $\mathcal{W}^{(0)}$, $z \to \infty$. For simplicity we limit the derivation to the Bethe lattice with nearest-neighbor hopping and in the end indicate how to generalize to other lattices. Thus $t_{\mathbf{R}\mathbf{R}'} = -\delta_{\langle \mathbf{R}\mathbf{R}' \rangle} t$ by Eq. (2.2), and in order to maintain finite kinetic energy t scales as $t \propto 1/\sqrt{z}$. This causes the source fields $(\eta^\dagger \eta)^n \propto t^{2n}$ in the nth order term to scale as $1/z^n$. Counteracting this is the number of terms in the sums over lattice sites $\mathbf{R}_n$ and $\mathbf{R}'_n$. But here the topology of the Bethe lattice simplifies the analysis. Removing one lattice site creates z disconnected lattices with no intermixing in their respective connected greens functions

$$G^{(0)}_{n;\mathbf{R}_1 ... \mathbf{R}_n, \mathbf{R}'_1 ... \mathbf{R}'_n}(\tau_1 ... \tau_n, \tau'_1 ... \tau'_n) = \delta_{\mathbf{R}_1, \mathbf{R}_2} ... \delta_{\mathbf{R}_1, \mathbf{R}_n} \delta_{\mathbf{R}_1, \mathbf{R}'_1} ... \delta_{\mathbf{R}_1, \mathbf{R}'_n} \times G^{(0)}_{n;\mathbf{R}_1 ... \mathbf{R}_1, \mathbf{R}_1 ... \mathbf{R}_1}(\tau_1 ... \tau_n, \tau'_1 ... \tau'_n) \,. \quad (4.19)$$

Thus the sum scales as z in all orders and the total scaling is $\frac{1}{z^{n-1}}$. In the limit $z \to \infty$, only the first-order term contributes and $\mathcal{W}^{(0)}$ is exactly given by

$$\lim_{z\to\infty} \mathcal{W}^{(0)}(\eta^\dagger,\eta) = -\int_0^\beta \int_0^\beta d\tau d\tau' \sum_\sigma c^\dagger_{0\sigma}(\tau) t^2 G^{(0)}(\tau - \tau') c_{0\sigma}(\tau') \quad (4.20)$$

where we have assumed translational invariance, $G^{(0)}_{\mathbf{R}\mathbf{R}} = G^{(0)}$.

This rather non-trivial result shows that the influence on a single site by the surroundings are governed by simple single-particle scattering without any of the many-body effects of the higher order Green's functions $G^{(0)}_{n;\mathbf{R}_1 ... \mathbf{R}_n, \mathbf{R}'_1 ... \mathbf{R}'_n}$ that otherwise contribute for finite z. Incorporating the

expression for $\mathcal{W}^{(0)}$ in the effective action S_{eff} now gives

$$\begin{aligned}
S_{\text{eff}} =& S_0 + \ln\langle e^{-\Delta S}\rangle_{S^{(0)}} + \ln \mathcal{Z}^{(0)} = S_0 + \lim_{z\to\infty} \mathcal{W}^{(0)}(\eta^\dagger,\eta) + \ln \mathcal{Z}^{(0)} = \\
&= \int_0^\beta d\tau \left(\sum_\sigma c^\dagger_{0\sigma}(\tau)(\partial_\tau - \mu) c_{0\sigma}(\tau) + U \hat{n}_{0\uparrow}(\tau)\hat{n}_{0\downarrow}(\tau) \right) \\
&\quad + \int_0^\beta \int_0^\beta d\tau d\tau' \sum_\sigma c^\dagger_{0\sigma}(\tau) t^2 G^{(0)}(\tau-\tau') c_{0\sigma}(\tau') + \ln \mathcal{Z}^{(0)} = \\
&= \int_0^\beta \int_0^\beta d\tau d\tau' \sum_\sigma c^\dagger_{0\sigma}(\tau) \left[\delta(\tau-\tau')(\partial_\tau - \mu) - t^2 G^{(0)}(\tau-\tau') \right] c_{0\sigma}(\tau') \\
&\quad + \int_0^\beta d\tau\, U \hat{n}_{0\uparrow}(\tau)\hat{n}_{0\downarrow}(\tau) + \ln \mathcal{Z}^{(0)}\,,
\end{aligned} \tag{4.21}$$

where the expression in brackets of the third last term can be identified as the inverse of the effective bath Green's function G_0 of the auxiliary impurity problem

$$G_0^{-1}(\tau-\tau') = \delta(\tau-\tau')(\partial_\tau - \mu) - t^2 G^{(0)}(\tau-\tau')\,. \tag{4.22}$$

The equation above only relates G_0 to the local connected Green's function $G^{(0)}$ of the lattice with cavity ($S^{(0)}$), but in the limit of infinite connectivity the creation of a cavity has negligible effect on the rest of the system and $G^{(0)}$ becomes equal to the local Green's function G_L of the full lattice system (S)

$$\begin{aligned}
\lim_{z\to\infty} G^{(0)}(\tau-\tau') = \lim_{z\to\infty} G^{(0)}_{\mathbf{R}\sigma}(\tau-\tau') = \lim_{z\to\infty} \langle c_{\mathbf{R}\sigma}(\tau) c^\dagger_{\mathbf{R}\sigma}(\tau')\rangle_{S^{(0)}} = \\
= \langle c_{\mathbf{R}\sigma}(\tau) c^\dagger_{\mathbf{R}\sigma}(\tau')\rangle_S = G_{L,\sigma}(\tau-\tau') = G_L(\tau-\tau')\,.
\end{aligned} \tag{4.23}$$

Thus replacing $G^{(0)}$ with G_L and transforming Eq. (4.22) to Matsubara frequencies gives the specific form of the self-consistency equation for the Bethe lattice

$$G_0^{-1}(i\omega_n) = i\omega_n + \mu - t^2 G_L(i\omega_n)\,. \tag{4.24}$$

To obtain the general self-consistency relation in Eq. (4.4) we have to take a slight detour and consider the non-interacting limit $U = 0$ of the above equation where we have that $G_0 = G_L$ and

$$\xi = i\omega_n + \mu = t^2 G_L + G_L^{-1} \;\Rightarrow\; G_L = \frac{\xi - \sqrt{\xi^2 - 4t^2}}{2t^2}\,,\ \mathrm{Im}[\xi] > 0\,. \tag{4.25}$$

As the non-interacting DOS $\rho^{(0)}(\epsilon)$ is the inverse Hilbert transform of the Green's function it can now be calculated

$$G_L(\xi) = \int d\epsilon \, \frac{\rho^{(0)}(\epsilon)}{\xi - \epsilon} \,, \quad \rho^{(0)}(\epsilon) = \frac{1}{2\pi t^2} \sqrt{\epsilon^2 - 4t^2} \,. \tag{4.26}$$

Returning to the interacting system $U \neq 0$, the interacting local Green's function G_L can also be calculated using $\rho^{(0)}(\epsilon)$ by including the local self-energy, $\Sigma(i\omega_n)$

$$G_L(\mathbf{R} = 0, i\omega_n) = \sum_{\mathbf{k}} G_L(\mathbf{k}, i\omega_n) = \int d\epsilon \, \frac{\rho^{(0)}(\epsilon)}{i\omega_n + \mu - \epsilon - \Sigma(i\omega_n)} \,, \tag{4.27}$$

and replacing ξ with $\tilde{\xi} = i\omega_n + \mu - \Sigma$, it has the same Hilbert transform as in Eq. (4.25) and the relation

$$i\omega_n + \mu - \Sigma = t^2 G_L + G_L^{-1} \,, \tag{4.28}$$

holds. Insertion in Eq. (4.24) finally gives the general self-consistency relation of Eq. (4.4) for the bath Green's function G_0 of the auxiliary impurity system

$$G_0^{-1} = G_L^{-1} + \Sigma \,. \tag{4.29}$$

The same relation can be shown to hold for more general lattices where the power counting in the coordination number z becomes more elaborate [17]. Note that Eq. (4.29) is not the Dyson equation for the full lattice Green's function $G_L(\mathbf{k}, i\omega_m)$, because only the *local* lattice Green's function $G_L(i\omega_n)$ of Eq. (4.3) enters.

4.2 Self-consistent equations

When brought together the equations for the lattice and impurity system forms a set of connected self-consistent equations. Given a lattice with infinite coordination and its non-interacting density of states $\rho^{(0)}(\epsilon)$, and an initial guess of the lattice self-energy $\Sigma_L(i\omega)$, it is possible to compute the corresponding impurity self-energy Σ_I through the following steps:

1. Calculate the local lattice Green's function G_L by integrating over the non-interacting density of states $\rho^{(0)}(\epsilon)$

$$G_L(i\omega_n) = \int d\omega \, \frac{\rho^{(0)}(\omega)}{i\omega_n - \omega + \mu - \Sigma_L(i\omega_n)} \,. \tag{4.30}$$

2. Determine the impurity bath Green's function G_0 using the self-consistency relation in Eq. (4.4) between the lattice and impurity

$$G_0^{-1} = G_L^{-1} + \Sigma_L \,. \tag{4.31}$$

3. Form the action S_{eff} of the auxiliary impurity problem using the bath Green's function G_0

$$\begin{aligned} S_{\text{eff}}[G_0] = & \int_0^\beta d\tau\, U \hat{n}_\uparrow(\tau)\hat{n}_\downarrow(\tau) - \\ & - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_\sigma c_\sigma^\dagger(\tau) G_0^{-1}(\tau - \tau') c_\sigma(\tau') \,, \end{aligned} \tag{4.32}$$

and calculate the local impurity Green's function

$$G_{I\sigma\sigma'}(i\omega_n) = \int_0^\beta d\tau\, e^{i\omega_n \tau} \langle c_\sigma(\tau) c_{\sigma'}^\dagger(0) \rangle_{S_{\text{eff}}[G_0]} \,. \tag{4.33}$$

4. Calculate the impurity self-energy Σ_I using the Dyson equation

$$\Sigma_I(i\omega_n) = G_0^{-1}(i\omega_n) - G_I^{-1}(i\omega_n) \,. \tag{4.34}$$

It is helpful to hide the details of the calculations by introducing a higher level of abstraction. To achieve this, the calculation steps 1 to 4 can be incorporated into a single function F that takes the lattice self-energy Σ_L as an argument, and returns the resulting impurity self-energy Σ_I

$$F_{U\beta}(\Sigma_L) = \Sigma_I \,, \tag{4.35}$$

where the external parameters of the calculation, the Hubbard coupling U and the inverse temperature β are indicated as subscripts of F. In terms of the function F the self-consistent DMFT solutions are nothing but the fix-point solutions

$$F_{U\beta}(\Sigma^*) = \Sigma^* \,, \quad \Sigma^* = \Sigma_L = \Sigma_I \,, \tag{4.36}$$

where the lattice and impurity self-energies coincide. The coupled equations and the fix-point function F is schematically shown in Fig. 4.3.

A closer inspection of the equations shows that the choice of the self-energy as fix-point variable is not unique. It is also possible to formulate the fix-point problem in terms of the bath Green's function G_0. This gives

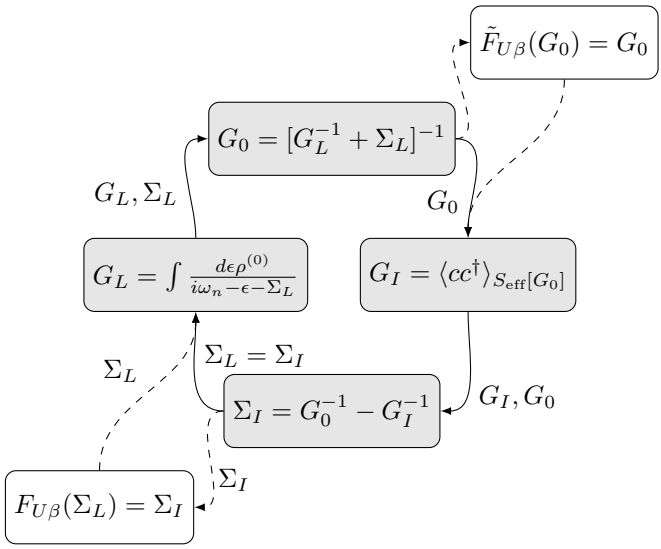


Figure 4.3: The self consistent DMFT equations and the two possible fix-point function formulations $F_{U\beta}(\Sigma_L)$ and $\tilde{F}_{U\beta}(G_0)$.

a function $\tilde{F}$ that takes the bath Green's function of the impurity as an argument and returns the bath Green's function generated by the lattice Green's function and self-energy

$$\tilde{F}_{U\beta}(G_0) = G_0 \,. \tag{4.37}$$

This parametrization of the fix-point function is also indicated in Fig. 4.3. In our calculations we have employed both formulations and they yield formally exactly the same solutions. Although it is often more beneficial to use one of the formulations depending on the particular method chosen to solve the impurity problem.

In later chapters we will study the properties of the fix-point function F in detail by explicitly calculating its Jacobian J_F. As F is a vector valued function, in the self-energy Σ at all Matsubara frequencies $\Sigma_n = \Sigma(i\omega_n)$, its Jacobian is formally given by

$$J_F(\Sigma) = \nabla_{\Sigma_m} F(\Sigma) = \frac{\partial F_n(\Sigma)}{\partial \Sigma_m} \,. \tag{4.38}$$

Unfortunately Σ_n is an infinite dimensional vector and the Jacobian J_F becomes an infinite dimensional matrix. To perform numerical calculations some finite dimensional (approximate) parametrization of Σ is needed, for example by mapping the impurity problem to a finite Anderson model or by simply using a finite number of Matsubara frequencies.

4.3 Fixpoint Solvers

In the previous section it was shown that the solutions to the DMFT self-consistent equations can be defined as fixpoints to the DMFT fixpoint function $F_{U\beta}$ of Eq. (4.36). It was also noted that there exists more than one possible choice of parametrization for $F_{U\beta}$, as indicated in Figure 4.3, using either the self-energy $\Sigma(i\omega_n)$ or the bath Green's function $G_0(i\omega_n)$. In next chapter yet another alternative parametrization, in terms of the finite Anderson Model parameters $\mathbf{x} = \{\epsilon_k, V_k\}$ [Eq. (5.28)], will be treated.

Although these mappings do differ in numerical aspects they all share leading order properties as they are re-parametrizations of the same fundamental DMFT mapping. In the following section we will discuss the general properties of $F_{U\beta}$ and adopt Σ as the fixpoint variable

$$F_{U\beta}(\Sigma) = \Sigma' \,, \tag{4.39}$$

keeping in mind that re-parametrizations in terms of G_0 and $\mathbf{x}$ are possible.

The task of finding fixpoints Σ^* to the fixpoint equation $F_{U\beta}(\Sigma^*) = \Sigma^*$ is a standard problem with large set of applicable numerical algorithms [63]. A common property of these algorithms is that they are local, in the sense that given some initial guess Σ_0 only the solution Σ^* "close" to the initial guess will be found. Thus finding all fixpoint solutions requires the use of several initial guesses as is the case for $F_{U\beta}$ in the vicinity of the Mott-transition.

In the DMFT literature the most commonly used fixpoint solver is forward recursion [17], but Newton methods has also received some attention [64]. Both algorithms experience numerical problems when applied at the Mott-transition. In this section we will review the merits and deficiencies of forward recursion and Newton methods giving special attention to the local stability in the vicinity of a fixpoint. The connection between the local stability properties and the spectral properties of the Jacobian J_F in Eq. (4.38) of the fixpoint function $F_{U\beta}$ will be shown. Finally, the understanding of the stability problems will be used to formulate the phase space extension method which is stable throughout the Mott-transition.

4.3.1 Forward recursion

The forward recursion algorithm is the most widely used fixpoint algorithm used in the DMFT community [17]. Its popularity can probably be traced back to the tradition of thinking in terms of the self-consistency cycle, shown in Fig. 4.3. The algorithm is a simple iterative evaluation of the DMFT fixpoint function $F_{U\beta}$. Given an initial guess Σ_0 the series $\{\Sigma_n\}$ is generated using the recursion relation

$$\Sigma_{n+1} = F_{U\beta}(\Sigma_n)\,, \tag{4.40}$$

and a fixpoint Σ^* is found if the series converges $\Sigma^* = \Sigma_\infty$. The convergence properties of the series in the vicinity of a fixpoint Σ^* can be assessed by Taylor-expanding $F_{U\beta}$ to first-order around Σ^*

$$F_{U\beta}(\Sigma^* + \delta\Sigma) \approx F_{U\beta}(\Sigma^*) + J_F(\Sigma^*) \cdot \delta\Sigma\,, \tag{4.41}$$

where $\delta\Sigma$ is a small perturbation. In the linear approximation of Eq. (4.41) the forward recursion series of a perturbed initial guess $\Sigma_0 = \Sigma^* + \delta\Sigma$ takes the form

$$\begin{aligned} \Sigma_0 =& \Sigma^* + \delta\Sigma\,, \\ \Sigma_n =& \Sigma^* + (J_F(\Sigma^*))^n \cdot \delta\Sigma\,. \end{aligned} \tag{4.42}$$

From this it is evident that the convergence is determined by the properties of the Jacobian at the fixpoint $J_F(\Sigma^*)$. The nth power of J_F is best described by expressing J_F in terms of its normalized eigenvectors $\mathbf{j}_i$ and eigenvalues ϵ_i

$$\begin{aligned} J_F &= \sum_i \epsilon_i\, \mathbf{j}_i \cdot (\mathbf{j}_i)^\dagger\,, \\ (J_F)^n &= \sum_i \epsilon_i^n\, \mathbf{j}_i \cdot (\mathbf{j}_i)^\dagger\,, \end{aligned} \tag{4.43}$$

where the orthonormality of $\mathbf{j}_i$ was used to obtain the last relation. Thus the convergence of Σ_n to Σ^* for a general perturbation $\delta\Sigma$ is determined by the eigenvalues ϵ_i of the Jacobian. To obtain convergence the powers of ϵ_i has to go to zero as $n \to \infty$, which requires that ϵ_i obey

$$|\epsilon_i| < 1\,, \forall i\,. \tag{4.44}$$

So forward recursion is only locally convergent if the spectral distribution of the fixpoint Jacobian $J_F(\Sigma^*)$ is bounded from above and below by one.

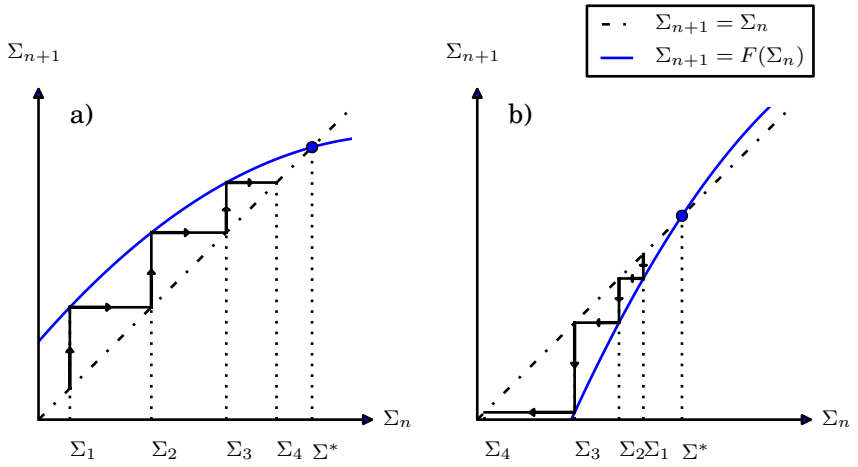


Figure 4.4: Forward recursion in one dimension for, a) the convergent case when $|\partial_\Sigma F_{U\beta}(\Sigma^*)| < 1$ and b) the divergent case when when $|\partial_\Sigma F_{U\beta}(\Sigma^*)| > 1$.

The speed of convergence of the forward recursion series $\{\Sigma_n\}$ is, due to Eq. (4.43), limited by the in magnitude largest eigenvalue ϵ as its nth power ϵ^n has the slowest decrease with respect to increasing n. This has the unpleasant effect that if $\epsilon \to 1^-$, when tuning the external parameters U and β the rate of convergence display a critical slowing down [65].

This very general statement can be more intuitively understood in one dimension where the Jacobian simplifies to the derivative of the fixpoint function $J_F(\Sigma) = \partial_\Sigma F_{U\beta}(\Sigma)$. In a two dimensional xy-plot of $F_{U\beta}$ as a function of Σ, $F_{U\beta}$ crosses the function $x = y$ at a fixpoint. In the same plot the forward recursion series can be generated by moving in zig-zag between these two functions, see Fig. 4.4. In the case when $|\partial_\Sigma F_{U\beta}(\Sigma^*)| < 1$ this causes the series $\{\Sigma_n\}$ to converge towards Σ^* while if $|\partial_\Sigma F_{U\beta}(\Sigma^*)| > 1$ even an initial guess Σ_0 very close to the fixpoint causes the series to diverge away from the fixpoint, see Fig. 4.4a and 4.4b respectively.

Thus when using forward recursion it is only possible to find the subset of all fixpoints that happens to have Jacobians with spectral distributions bounded by one in magnitude. Even though this is a large limitation of the accessible fixpoint space, forward recursion has the great advantage of being extremely simple to implement and to converge fast when the interesting fixpoints have small Jacobian eigenvalues $|\epsilon_i| \ll 1$.

4.3.2 Newton methods

The family of Newton methods are actually multidimensional root solvers but can easily be applied to find fixpoints of a function $F_{U\beta}$ by finding roots of the associated root function $R_{U\beta}$ defined by

$$R_{U\beta}(\Sigma^*) \equiv F_{U\beta}(\Sigma^*) - \Sigma^* = \mathbf{0}\,. \tag{4.45}$$

The original Newton's method [63] then generates the series

$$\Sigma_{n+1} = \Sigma_n - (J_R(\Sigma_n))^{-1} R_{U\beta}(\Sigma_n)\,, \tag{4.46}$$

where J_R is the Jacobian of the root function $R_{U\beta}$. The required explicit computation of J_R at each iteration makes this approach ill suited for high-dimensional problems, as a numerical approximation of J_R requires as many function evaluations as the dimension of $R_{U\beta}$. To overcome this limitation a zoo of algorithms has been developed that approximates the Jacobian using information from previous iterations. In this work Broyden's second method [66], which approximate the Jacobian by rank-one updates, has been used. The application of Broyden's method directly to the DMFT root function $R_{U\beta}$ has already been explored in Ref. [64].

To assess the local stability of Newton's method in the vicinity of a fixpoint Σ^* we again apply a perturbation $\delta\Sigma$ sufficiently small to stay in the linear regime where $R_{U\beta}(\Sigma^*+\delta\Sigma) \approx J_R(\Sigma^*)\cdot\delta\Sigma$ and $J_R(\Sigma^*+\delta\Sigma) \approx J_R(\Sigma^*)$. Now the series Σ_n converges in one iteration when given the perturbed initial guess

$$\begin{aligned} \Sigma_0 =& \Sigma^* + \delta\Sigma\,, \\ \Sigma_1 =& \Sigma_0 - J_R^{-1}(\Sigma_0) R_{U\beta}(\Sigma_0) \approx \Sigma^*\,, \end{aligned} \tag{4.47}$$

where we have assumed that the Jacobian $J_R(\Sigma^*)$ is invertible (non-singular). Thus Newton's method is locally convergent as long as $J_R(\Sigma^*)$ has no zero eigenvalues. To compare with forward recursion J_R is connected to J_F by

$$J_R(\Sigma) = \nabla R_{U\beta}(\Sigma) = \nabla F_{U\beta}(\Sigma) - \mathbf{1} = J_F(\Sigma) - \mathbf{1}\,, \tag{4.48}$$

so in terms of the eigenvalues ϵ_i of the Jacobian J_F of the DMFT function $F_{U\beta}$ Newton's method is stable when

$$\epsilon_i \neq 1\,, \forall i\,. \tag{4.49}$$

This requirement is a large improvement compared to forward recursion and Eq. (4.44), the respective ranges of stability are sketched in Fig. 4.5. Using

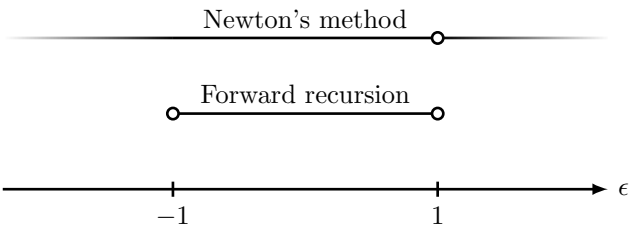


Figure 4.5: Local stability regions of Newton's method and forward recursion in terms of the spectral distribution of eigenvalues of the Jacobian J_F of the DMFT fixpoint function $F_{U\beta}$ at a fixpoint Σ^*. Open circles denotes missing points on the intervals.

one of the modified Newton methods, approximating the Jacobian, the gain in stability comes with only a slight increase in computational cost. The less restrictive local stability requirements of Newton methods make is possible to find a larger set of all fixpoints of $F_{U\beta}$. In particular it is possible to locate fixpoints that are impossible to find using forward recursion.

4.3.3 Phase Space Extension

As shown in Paper I the Jacobian J_F of the DMFT fixpoints becomes singular at the hysteresis boundaries of the Mott-transition, causing both forward recursion and Newton's method to fail. This behavior is highly undesirable when studying the development of the fixpoints as a function of the external parameters, the Hubbard U and the inverse temperature β, in the vicinity of the Mott-transition. To counter this we have devised a simple phase space extension algorithm that can continuously map out the DMFT fixpoints throughout the entire hysteresis region.

The phase space extension is based on the idea of using the external parameters as additional degrees of freedom in the search for fixpoint solutions. By viewing U as an external field it is possible to find a scalar conjugate field $A(\Sigma)$ having different values for the full set of coexisting fixpoints. This property causes A to have non-zero derivative with respect to the fixpoint parametrization Σ, $\nabla_\Sigma A(\Sigma) \neq 0$, at the points where J_F becomes singular. Now given the scalar function $A(\Sigma)$ the extended root problem $\tilde{R}_{\alpha\beta}$ is constructed as

$$\tilde{R}_{\alpha\beta}(U, \Sigma) = (\alpha - A(\Sigma), R_{U\beta}(\Sigma)) = \mathbf{0}\,, \tag{4.50}$$

where U is a free parameter and the value of A is restricted to the new

external parameter α by requiring that the first entry in $\tilde{R}_{\alpha\beta}$ is zero, $\alpha - A(\Sigma) = 0$. The fact that A lifts the degeneracy of coexisting fixpoints causes the Jacobian $J_{\tilde{R}}$ of the extended root problem to be invertible even when J_R is singular. Thus when sweeping α, all fixpoints of $F_{U\beta}$ along an isotherm can be determined by solving Eq. (4.50) using a (quasi) Newton method.

In the case of the Mott-transition of the Hubbard model and the coupling strength U, one of the possible choices of conjugate the field is the double occupancy $D(\Sigma)$. But in the numerical studies of this thesis the value of the self-energy at the first Matsubara frequency, $A(\Sigma) \equiv \Sigma'(i\omega_0) = F_{U\beta}(\Sigma)(i\omega_0)$, was used. This choice is motivated by $\Sigma'(i\omega_0)$ being one-to-one to the Eliashberg estimate of the quasi-particle weight, see Ref. [67] for details.

Example

The reason why the phase space extension gives an invertible Jacobian even in the cases when the original root problem does not can be understood when applied to a simple solvable model. In Section 2.6 a phenomenological simplified Landau function $\mathcal{L}$ was introduced and defined as

$$\mathcal{L}(u,t,d) = \beta_0 ud + \gamma_1 t\frac{d^2}{2} + \frac{d^4}{4}\,, \tag{4.51}$$

where β_0 and γ_1 are model parameters. The analog of the DMFT fixpoints are the stationary points of the Landau function and the surface $S_{\mathcal{L}}$

$$S_{\mathcal{L}} = \{(u,t,d) \,:\, \partial_d \mathcal{L}(u,t,d) = \beta_0 u + \gamma_1 td + d^3 = 0\}\,. \tag{4.52}$$

This can be put in the form of a fixpoint problem $F_{ut}(d)$ by rearranging the terms in the stationary point equation

$$F_{ut}(d) = -\frac{d^3 + \beta_0 u}{\gamma_1 t}\,, \tag{4.53}$$

where the points on $(u,t,d^*) \in S_{\mathcal{L}}$ are fixpoint solutions, $F_{ut}(d^*) = d^*$. To formulate the phase space extension, also the root function R_{ut} is needed and the definition in Eq. (4.45) gives

$$R_{ut}(d) = F_{ut}(d) - d = -\frac{1}{\gamma_1 t}\left(\beta_0 u + \gamma_1 td + d^3\right) = 0\,. \tag{4.54}$$

The scalar function conjugate to the coupling u can in this case be chosen as the double occupancy $d' = F_{ut}(d)$, given by the fixpoint function. The extended root problem $\tilde{R}_{\tilde{d}t}(u, d)$ is then given by

$$\tilde{R}_{\tilde{d}t}(u,d) = (\tilde{d} - F_{ut}(d), R_{ut}(d)) = (\tilde{d} + \frac{d^3 + \beta_0 u}{\gamma_1 t}, -\frac{1}{\gamma_1 t}\left(\beta_0 u + \gamma_1 t d + d^3\right)). \tag{4.55}$$

Knowing $\tilde{R}$ analytically enables a direct calculation of its Jacobian

$$J_{\tilde{R}} = \nabla \tilde{R}_{\tilde{d}t}(u,d) = \frac{1}{\gamma_1 t}\begin{pmatrix} \beta_0 & 3d^2 \\ -\beta_0 & -(\gamma_1 t + 3d^2) \end{pmatrix}, \tag{4.56}$$

having the eigenvalues

$$\lambda = -\frac{1}{2}(-\beta_0 + \gamma_1 t + 3d^2) \pm \sqrt{\frac{1}{4}(-\beta_0 + \gamma_1 t + 3d^2)^2 + \beta_0 \gamma_1 t}\,. \tag{4.57}$$

Now the points where the Jacobian J_R of the root function $R_{ut}(d)$ becomes singular are the saddle-node bifurcation lines as discussed in Section 2.6. For our simple Landau function J_R is scalar and equal to zero when singular

$$J_R = -1 - \frac{3d^2}{\gamma_1 t} = 0 \quad \Rightarrow \quad \gamma_1 t + 3d^2 = 0\,. \tag{4.58}$$

Insertion into the equation for the eigenvalues λ of the extended root problem $\tilde{R}$ gives

$$\lambda = \frac{\beta_0}{2} \pm \sqrt{\frac{\beta_0^2}{4} + \beta_0 \gamma_1 t}\,. \tag{4.59}$$

As $\beta_0, \gamma_0 > 0$ criticality occur when $t \lesssim 0$, which makes the eigenvalues λ real and non-zero and ensures that $J_{\tilde{R}}$ is invertible. In other words a Newton's method applied to $\tilde{R}_{\tilde{d}t}(u, d)$ converges locally for the points where the Jacobian of $R_{ut}(d)$ is not invertible.

Chapter 5

Auxiliary Impurity Problem

The mapping of the lattice system to an auxiliary impurity system in DMFT was described in Chapter 4, where also the self-consistency condition determining the impurity bath Green's function G_0 was derived. Combined with the local interaction of the lattice, G_0 fully specifies the effective action S_{eff} of the impurity

$$S_{\text{eff}}[G_0] = \int_0^\beta d\tau\, U\hat{n}_\uparrow(\tau)\hat{n}_\downarrow(\tau) - \int_0^\beta d\tau \int_0^\beta d\tau' \sum_\sigma c_\sigma^\dagger(\tau) G_0^{-1}(\tau-\tau') c_\sigma(\tau')\,. \quad (5.1)$$

In order to close the self-consistent DMFT loop of Fig. 4.3 the local interacting Green's function G_I of the impurity has to be calculated

$$G_{I\sigma\sigma'}(i\omega_n) = \int_0^\beta d\tau\, e^{i\omega_n\tau} \langle c_\sigma(\tau) c_{\sigma'}^\dagger(0)\rangle_{S_{\text{eff}}[G_0]}\,. \quad (5.2)$$

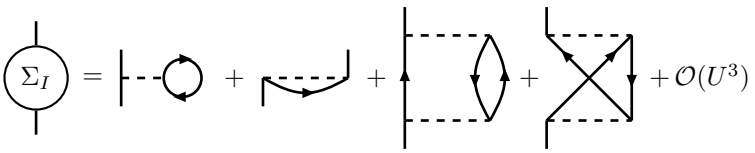
This is of course still a formidable task and much research has been put into developing analytical and numerical methods for this particular problem. In the literature this family of methods are commonly referred to as impurity solvers. For an overview of methods see Refs. [17, 68, 69].

I this thesis we will only treat the impurity solvers Iterated Perturbation Theory (IPT), Exact Diagonalization (ED), and the Distributional Exact

Diagonalization extension (Dist-ED) of ED that we have developed. In the following sections we will introduce IPT, discuss ED in more detail, and give a background to Dist-ED that is treated in more detail in Paper II.

5.1 Iterated Perturbation Theory

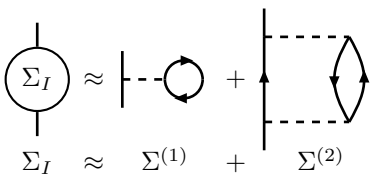
The approach of Iterated Perturbation Theory [70–72] is based on a simple truncation of the diagrammatic expansion of the impurity self-energy Σ_I to second order in U, giving the four diagrams



. (5.3)

The effective action of the impurity in Eq. (5.1), is actually mappable to an Anderson impurity model [73] for which the bath Green's function G_0 plays the role of the free propagator. The next section is devoted to the details of the mapping. Here we just note that perturbation theory applied to the Anderson model were studied before the dawn of DMFT [74–77].

In the case of the Hubbard model the resulting DMFT impurity model only contains a density-density interaction, $U\hat{n}_\uparrow\hat{n}_\downarrow$, which requires different spin labels for the propagators on each side of an interaction vertex. This symmetry is only fulfilled by the Hartree and bubble diagram in the general expansion while the others give zero contribution. Thus approximating Σ_I to second order requires the calculation of

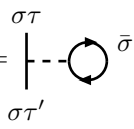


. (5.4)

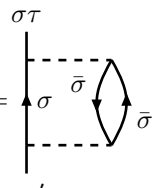
The standard rules for unlabeled Feynman diagrams [26] give algebraic ex-

pressions for the first- and second-order contributions $\Sigma^{(1)}$ and $\Sigma^{(2)}$ as

$$\Sigma^{(1)}_\sigma(\tau-\tau') = \quad = -\delta(\tau-\tau')UG_{0\bar{\sigma}}(0^-)\,,$$



$$\Sigma^{(2)}_\sigma(\tau-\tau') = \quad = -U^2G_{0\sigma}(\tau-\tau')G_{0\bar{\sigma}}(\tau-\tau')G_{0\bar{\sigma}}(\tau'-\tau)\,. \tag{5.5}$$



Now assuming a paramagnetic solution where the bath Green's function G_0 is independent on spin $G_{0,\uparrow}(i\omega_n) = G_{0,\downarrow}(i\omega_n)$, and transforming G_0 from imaginary frequencies $i\omega_n$ to imaginary time τ gives

$$\begin{aligned}
G_0(\tau) &= \frac{1}{\beta}\sum_n G_0(i\omega_n)e^{-i\omega_n\tau} \\
\Sigma^{(1)}_I(i\omega_n) &= -UG_0(0)\int_0^\beta d\tau\,\delta(\tau)e^{i\omega_n\tau} = -U\langle\hat{n}\rangle \\
\Sigma^{(2)}_I(i\omega_n) &= -U^2\int_0^\beta d\tau\,G_0^2(\tau)G_0(-\tau)e^{i\omega_n\tau}\,,
\end{aligned}$$

where we have used that $G_0(0) = \langle\hat{n}\rangle$. As the first-order Hartree diagram $\Sigma^{(1)}_I$ only give a constant contribution it can be absorbed in to the chemical potential $\tilde{\mu} = \mu - \langle\hat{n}\rangle U$. Thus, the numerical implementation of IPT mainly consists of a Fourier transformation $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ and can be summarized as

$$\begin{aligned}
G_0(\tau) &= \mathcal{F}^{-1}\left\{G_0(i\omega_n)\right\}\,, \\
\Sigma_I(i\omega_n) &= -U^2\mathcal{F}\left\{G_0^2(\tau)G_0(-\tau)\right\}\,.
\end{aligned}$$

An important question to ask is why IPT is relevant at all for the description of the Hubbard model for any sizable electron-electron coupling U. Being a second order approximation one could expect it to work in some range of weak coupling where higher order diagrams does not contribute. To be able to describe both weak any strong correlation physics any impurity solver also have to contain the atomic limit.

In the case of IPT some further investigation of this limit is required. Taking the atomic limit causes the electron dispersion to become momentum independent $\epsilon_{\mathbf{k}} = t$ and the non-interacting Green's function takes the form

$$G_{\mathrm{atm},0} = \frac{1}{i\omega_n - t + \mu} \tag{5.6}$$

and the analytical IPT approximation to the self-energy can be evaluated to

$$\Sigma^{\mathrm{IPT}}_{\mathrm{atm},\sigma}[G_{\mathrm{atm},0}] = \Sigma^{(1)}_{\sigma} + \Sigma^{(2)}_{\sigma} = \langle \hat{n}_{\bar{\sigma}} \rangle U + \frac{\langle \hat{n}_{\bar{\sigma}} \rangle (1 - \langle \hat{n}_{\bar{\sigma}} \rangle) U^2}{i\omega_n - t + \mu - \langle \hat{n}_{\bar{\sigma}} \rangle U}\,, \tag{5.7}$$

if one includes the Hartree term in the chemical potential. Comparing with the self-energy for the real atomic system in Eq. (2.26) the only difference is the prefactor for the U term in the denominator. But by a lucky coincidence these two prefactors coincide at half-filling and because of this IPT also has the correct atomic limit at half-filling.

IPT was one of the first impurity solvers applied to DMFT and for the half-filled single-band Hubbard model it has been particularly important for understanding the Mott-Hubbard metal insulator transition [70–72]. It has also acted as a reference for the development of more sophisticated methods like Quantum Monte Carlo solvers [67]. On the downside IPT is limited to the single-band case at half-filling in its original formulation. However there has been some attempts to perform the same type of non-interacting to atomic limit interpolation for arbitrary fillings [78].

5.2 Anderson model mapping

The Anderson model [73] was initially proposed by P. W. Anderson in 1961 as an effective model for magnetic impurities in metallic systems. Since then a plethora of approximate methods has been developed to handle the model, one notable example is the renormalization group method [79]. The basic ingredients in the single-band version is a single impurity site with local electron-electron interaction and a surrounding itinerant non-interacting electron system with dispersion $\epsilon_{\mathbf{k}}$. The coupling of the impurity to its surroundings is described by a hybridization $V_{\mathbf{k}}$ with the eigenstates of the non-interacting electrons. In second quantization the Hamiltonian takes the

form

$$\hat{H}_{\text{AM}} = \sum_{\sigma}(\epsilon_I - \mu)c_\sigma^\dagger c_\sigma + U\hat{n}_\uparrow\hat{n}_\downarrow + \sum_{k=1}^{\infty}\sum_{\sigma}\left[\epsilon_k c_{k\sigma}^\dagger c_{k\sigma} + V_k\left(c_\sigma^\dagger c_{k\sigma} + c_{k\sigma}^\dagger c_\sigma\right)\right], \quad (5.8)$$

where $c_\sigma^\dagger$ and c_σ are impurity creation and annihilation operators, $c_{k\sigma}^\dagger$ and $c_{k\sigma}$ are the equivalent operators for the non-interacting system and the momentum $\mathbf{k}$ has been replaced with a general quantum label k.

The connection between the Anderson model and the auxiliary impurity problem can be established using the philosophy of the cavity construction derivation of DMFT. The local electron-electron interaction is the same in both cases but the representability of the general bath Green's function $G_0(i\omega_n)$ by the non-interacting electron system in the Anderson model is not straight forward. A division of the action of the Anderson model in an impurity S_0, a non-interacting surrounding $S^{(0)}$ and a hybridization term ΔS gives

$$\begin{aligned}
S_0 &= \int_0^\beta d\tau\left[\sum_\sigma c_\sigma^\dagger(\tau)(\partial_\tau + \epsilon_I - \mu)c_\sigma(\tau) + U\hat{n}_\uparrow(\tau)\hat{n}_\downarrow(\tau)\right], \\
S^{(0)} &= \int_0^\beta d\tau\sum_{k=1}^{\infty}\sum_\sigma c_{k\sigma}^\dagger(\tau)(\partial_\tau + \epsilon_k - \mu)c_{k\sigma}(\tau), \\
\Delta S &= \int_0^\beta d\tau\sum_{k=1}^{\infty}\sum_\sigma V_k(c_{k\sigma}^\dagger(\tau)c_\sigma(\tau) + c_\sigma^\dagger(\tau)c_{k\sigma}(\tau)) = \\
&= \int_0^\beta d\tau\sum_{k=1}^{\infty}\sum_\sigma (c_{k\sigma}^\dagger(\tau)\eta_{k\sigma}(\tau) + \eta_{k\sigma}^\dagger(\tau)c_{k\sigma}(\tau)), \qquad (5.9)
\end{aligned}$$

where the source terms η are defined as, $\eta_{k\sigma} = V_k c_\sigma$. Integrating out the $c_{k\sigma}$ and $c_{k\sigma}^\dagger$ fields gives again the generating functional $\mathcal{W}^{(0)}(\eta^\dagger, \eta)$ for connected Green's functions of the non-interacting electron system. The non-interacting property causes only the single particle connected Green's function to be non-zero so only the first term in the perturbation expansion of $\mathcal{W}^{(0)}(\eta^\dagger, \eta)$ contribute to the effective local action of the Anderson impurity

$$\begin{aligned}
\mathcal{W}^{(0)}(\eta^\dagger, \eta) &= \ln\langle e^{-\Delta S}\rangle_{S^{(0)}} \\
&= \sum_{k=1}^{\infty}\sum_\sigma c_\sigma^\dagger\frac{|V_k|^2}{i\omega_n - \epsilon_k}c_\sigma = \sum_\sigma c_\sigma^\dagger\Delta(i\omega_n)c_\sigma, \quad (5.10)
\end{aligned}$$

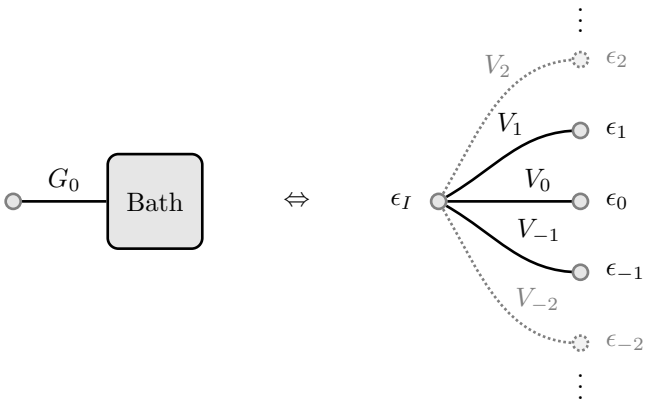


Figure 5.1: Schematic mapping between the auxiliary impurity problem (left) to the Anderson model (right).

where we have introduced the hybridization function $\Delta(i\omega_n)$

$$\Delta(i\omega_n) = \sum_{k=1}^{\infty} \frac{|V_k|^2}{i\omega_n - \epsilon_k} \,. \tag{5.11}$$

The effective action S_{eff} for the Anderson impurity is then given by

$$\begin{aligned} S_{\text{eff}} = &\int_0^\beta d\tau \, U \hat{n}_\uparrow(\tau) \hat{n}_\downarrow(\tau) \\ &+ \int_0^\beta \int_0^\beta d\tau d\tau' \sum_\sigma c_\sigma^\dagger(\tau) \left[\delta(\tau - \tau')(\partial_\tau + \epsilon_I - \mu) - \Delta(\tau - \tau') \right] c_\sigma(\tau') \,. \end{aligned} \tag{5.12}$$

In terms of the hybridization function $\Delta(i\omega_n)$ the final relation between the auxiliary impurity bath Green's function and the Anderson model formulation after transforming to Matsubara frequency becomes

$$G_0^{-1}(i\omega_n) = i\omega_n + \mu - \epsilon_I - \Delta(i\omega_n) \,, \tag{5.13}$$

where the representability of $G_0(i\omega_n)$ is guaranteed by the infinite number of poles in $\Delta(i\omega_n)$, for a schematic view of the mapping see Fig. 5.1. Although the mapping of the auxiliary impurity model to an Anderson model is straight forward once the theoretical framework of DMFT is established, it has been important for the rapid development of the field. When

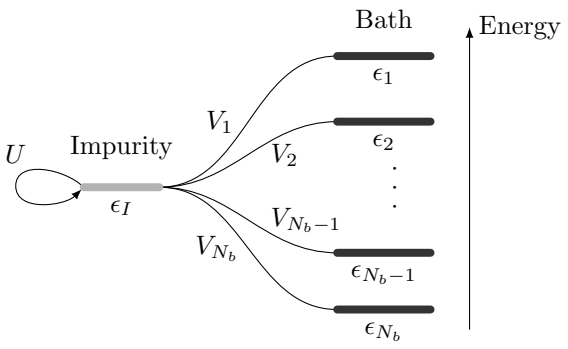


Figure 5.2: Finite single-impurity Anderson model, with one correlated impurity level ϵ_I, bath levels ϵ_k, $1 \leq k \leq N_b$, hybridizations V_k and Hubbard repulsion U.

DMFT was formulated in the beginning of the 90s, numerical methods for the Anderson model had already been under development for 30 years. The emergence of DMFT has ever since been a driving force for the development of more powerful methods.

5.3 Exact Diagonalization

The second impurity solver used in this thesis is Exact Diagonalization (ED) introduced by Caffarel and Krauth [80]. As the name indicates ED is based on the exact solution of a finite model by finding all eigenstates through diagonalization of its matrix-representation in some basis. The auxiliary impurity problem is first mapped to the infinite dimensional Anderson model which is then truncated to a finite number of sites. By limiting the summation over k in Eq. (5.8) to some number N_b, this gives the truncated finite single-impurity Anderson model used in ED with the Hamiltonian

$$\hat{H}_{\text{ED}} = \sum_{\sigma}(\epsilon_I - \mu)c^{\dagger}_{\sigma}c_{\sigma} + U\hat{n}_{\uparrow}\hat{n}_{\downarrow} + \sum_{k=1}^{N_b}\sum_{\sigma}\epsilon_k c^{\dagger}_{k\sigma}c_{k\sigma} + \sum_{k=1}^{N_b}\sum_{\sigma}V_k\left(c^{\dagger}_{\sigma}c_{k\sigma} + c^{\dagger}_{k\sigma}c_{\sigma}\right). \quad (5.14)$$

see Fig. 5.2 for a schematic picture. With a finite set of states $\{k\sigma\}$ the Hilbert space of $\hat{H}_{\text{ED}}$ is finite and in principle it is possible to construct its matrix representation, determine all eigenstates and calculate the interacting impurity Green's function $G_I(i\omega_n)$.

The gain of getting an exactly solvable model comes at the price of approximating the impurity bath Green's function $G_0(i\omega_n)$, having an infinite set of poles, with a bath Green's function $G_{0,\mathrm{ED}}(i\omega_n)$ having only a finite number of poles. Using the exact impurity hybridization function $\Delta(i\omega_n)$ and Eqs. (5.13) and (5.11) this approximation can be written as

$$\begin{aligned} G_0^{-1}(i\omega_n) &= i\omega_n + \mu - \epsilon_I - \Delta(i\omega_n) \\ &\approx i\omega_n + \mu - \epsilon_I - \Delta_{\mathrm{ED}}(i\omega_n) = G_{0,\mathrm{ED}}^{-1}(i\omega_n)\,, \\ \Delta_{\mathrm{ED}}(i\omega_n) &= \sum_{k=1}^{N_b} \frac{|V_k|^2}{i\omega_n - \epsilon_k}\,, \end{aligned} \tag{5.15}$$

where the energies ϵ_k and hybridization strengths V_k of the electronic states with labels $\{k\sigma\}$ are called bath states as they determine the bath Green's function $G_{0,\mathrm{ED}}$.

5.3.1 Bath mapping

In practical DMFT calculations the bath Green's function G_0 is obtained numerically and the corresponding finite model is determined by fitting the bath Green's function $G_{0,\mathrm{ED}}$ of the finite system to G_0 by treating ϵ_k and V_k as fitting parameters. Although this minimization only has a small number of free parameters the Green's functions are complex vectors in the infinite-dimensional vector-space of the Matsubara frequencies $i\omega_n$. On top of that there is no natural norm in this vector space that can be used to measure the distance between two vectors, Fig. 5.3 gives a schematic picture of the situation.

The issue of the infinite number of Matsubara frequencies are usually solved by disregarding all frequencies above some cut-off frequency ω_{N_c}. This is possible since the analytical structure of Matsubara Green's functions gives the asymptotic behavior $\sim \frac{1}{i\omega_n}$. Setting N_c large enough then only discards the known asymptotic tail.

Regarding what distance measure χ^2 to use in the now finite dimensional vector space, Caffarel and Krauth [80] proposed that the absolute square sum of the difference between the inverse of the two bath Green's functions could be used

$$\chi^2_{\mathrm{CK}}(\epsilon_k, V_k) = \frac{1}{N_c + 1} \sum_{n=0}^{N_c} \left| G_0^{-1}(i\omega_n) - (G_{0,\mathrm{ED}}^{(\epsilon_k, V_k)})^{-1}(i\omega_n) \right|^2 . \tag{5.16}$$

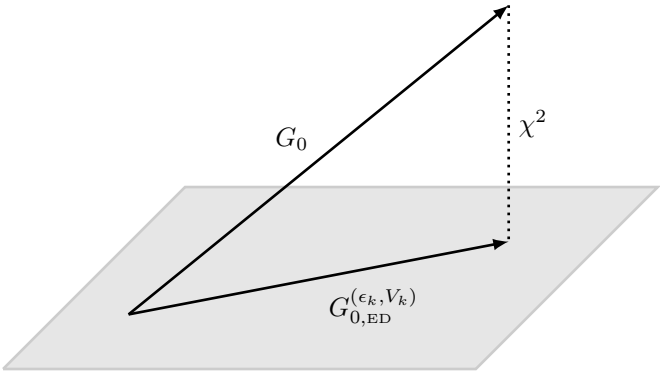


Figure 5.3: Schematic picture of the projection of the impurity bath Green's function G_0 on to the hyper-surface of finite Anderson Model representable ED bath Green's functions $G_{0,\mathrm{ED}}$ by minimization of the distance function χ^2.

Note that χ^2 sometimes is referred to as the "penalty function" as it has to be minimized in ED. Since the proposal of Eq. (5.16) many authors have investigated numerous other forms of the distance function χ^2 [81–84]. One of the problems with the original form of Eq. (5.16) is that it does not discriminate between low and high frequencies. This causes the particular choice of N_c to be important, and when set too high the fit is only sensitive to the asymptotic tail. One proposed way to counter this behavior is to weight the Green's function differences with some negative power of the Matsubara frequencies. In Paper I we used the distance function

$$\chi^2(\epsilon_k, V_k) = \frac{1}{N_c+1} \sum_{n=0}^{N_c} \left| G_0(i\omega_n) - G_{0,\mathrm{ED}}^{(\epsilon_k,V_k)}(i\omega_n) \right|^2 , \tag{5.17}$$

and found when comparing it to the forms based on differences between inverses of the Green's functions that it is more stable when setting a high N_c. To compare it with the original form of Eq. (5.16) it can be rewritten as

$$\chi^2[\epsilon_k, V_k] = \frac{1}{N_c+1} \sum_{n=0}^{N_c} \left| G_0 G_{0,\mathrm{ED}} \right| \left| G_0^{-1} - G_{0,\mathrm{ED}}^{-1} \right|^2 , \tag{5.18}$$

showing that its high frequency behavior is equal to Eq. (5.16) with an extra $1/\omega_n^2$ factor damping out the weight of the asymptotic tail.

The arbitrariness in the choice of χ^2 is generally seen as one of the weaknesses of ED as different choices potentially could produce different results. Calculations using the distance functions of both Eq. (5.16) and Eq. (5.17) indicate that the properties of the critical second-order end-point of the metal-insulator transition is insensitive to this choice. Given a distance function χ^2 the free parameters ϵ_k and V_k are readily determined using a standard conjugate gradient minimization algorithm, which fully determines the ED Hamiltonian $\hat{H}_{\text{ED}}$.

5.3.2 Diagonalization

Given the Hamiltonian $\hat{H}_{\text{ED}}$ in second quantization form, its corresponding matrix representation can only be found relative to some many-body basis $\{\Gamma\}$. A convenient choice is the occupation-number basis containing every possible combination of local occupations

$$|\Gamma\rangle = |n_\uparrow n_{1\uparrow} ... n_{N_b\uparrow}; n_\downarrow n_{1\downarrow} ... n_{N_b\downarrow}\rangle\,, \tag{5.19}$$

where $n_\sigma = 0, 1$ are the occupations of the impurity, and $n_{k\sigma} = 0, 1$ the occupations of the bath sites. The representations of the creation and annihilation operators $c_i^\dagger$, c_i can be determined by their action on a general occupation-number state [14]

$$\begin{aligned} c_i^\dagger|n_1 \dots n_{2N_b+2}\rangle = &\quad (1-n_i)(-1)^{\sum_{j<i} n_j}|n_1 \dots (n_i+1) \dots n_{2N_b+2}\rangle\,, \\ c_i|n_1 \dots n_{2N_b+2}\rangle = &\quad n_i(-1)^{\sum_{j<i} n_j}|n_1 \dots (n_i-1) \dots n_{2N_b+2}\rangle\,, \end{aligned} \tag{5.20}$$

where i is a super index of all site and spin labels $1 < i < 2N_f$, where N_f is the number of electron sites $N_f = N_b + 1$. Knowing the particular representations of the creation operator $(c^\dagger)_{\Gamma\Gamma'}$ and annihilation operator $(c)_{\Gamma\Gamma'}$ it is straight forward to construct $(\hat{H}_{\text{ED}})_{\Gamma\Gamma'}$ [85]. Numerical diagonalization can now be used to calculate all eigenvalues E_ν and eigen-states $|\nu\rangle$ of $\hat{H}_{\text{ED}}$, $\hat{H}_{\text{ED}}|\nu\rangle = E_\nu|\nu\rangle$.

With access to the full set of eigenstates the impurity Green's function G_I is given by the Lehman spectral representation [14, 86]

$$G_{I\sigma}(i\omega_n) = \frac{1}{\mathcal{Z}} \sum_{\nu\mu} \frac{|\langle\mu|c_\sigma^\dagger|\nu\rangle|^2}{i\omega_n + E_\nu - E_\mu} \left(e^{-\beta E_\nu} + e^{-\beta E_\mu}\right)\,, \tag{5.21}$$

where the grand partition function $\mathcal{Z}$ of the impurity is calculated by the

trace over eigenstates

$$\mathcal{Z} = \mathrm{Tr}\left[e^{-\beta \hat{H}_{\mathrm{ED}}}\right] = \sum_{\nu} e^{-\beta E_\nu} \,. \tag{5.22}$$

As the local Green's function of the impurity and lattice coincide for self-consistent DMFT solutions local properties of the lattice coincide with the impurity. Such an important local quantity is the on-site double occupancy D of Eq. (2.12) that is obtained as

$$D = \frac{\langle \hat{D} \rangle}{\mathcal{N}} = \sum_{\nu} \langle \nu | \hat{n}_\uparrow \hat{n}_\downarrow | \nu \rangle e^{-\beta E_\nu} \,. \tag{5.23}$$

The full diagonalization and Lehman representation are both straight forward to implement but both operations scales as N_Γ^3 with respect to the size of the Hilbert space N_Γ. This scaling is highly relevant as the size of the Hilbert space scales exponentially with the number of electronic states N_f giving $N_\Gamma = 2^{2N_F}$.

To reduce the dimensions of the diagonalization problem it is possible block-diagonalize the Hamiltonian by using its symmetry properties [85]. The single-band Anderson model Hamiltonian conserves both particle number, magnetization, and total spin, as can be confirmed by evaluating the commutations

$$[\hat{H}_{\mathrm{ED}}, \hat{N}] = 0 \,, \quad [\hat{H}_{\mathrm{ED}}, \hat{S}_z] = 0 \,, \quad [\hat{H}_{\mathrm{ED}}, \hat{\mathbf{S}}^2] = 0 \,. \tag{5.24}$$

The two first symmetries can be equivalently expressed in terms of the total number of spin up and spin down $\hat{N} = \hat{N}_\uparrow + \hat{N}_\downarrow$, $\hat{S}_z = \hat{N}_\uparrow - \hat{N}_\downarrow$. Thus grouping states Γ with the same number of spin up and down states $(N_\uparrow, N_\downarrow)$, block-diagonalizes the Hamiltonian in blocks with dimensions $\binom{N_f}{N_\uparrow}\binom{N_f}{N_\downarrow}$. The rotational invariance can be used to further reduce this but requires actual diagonalization, and does not correspond to a simple regrouping of the occupation-number states, as they are not eigenstates of $\hat{\mathbf{S}}^2$. As an example of the gain in block diagonalization the case of $N_f = 6$ gives a Hilbert space with the dimension $N_\Gamma = 4096$, that after block diagonalization consist of 12 blocks where the largest one is $(N_\uparrow, N_\downarrow) = (3, 3)$ with the dimension 400.

An interesting further development is the Krylov subspace methods [87] that can be used to diagonalize even larger matrices requiring only the matrix vector multiplication operation. These methods have also been extended to the calculation of finite temperature Green's functions through a two step method [86, 88, 89] enabling the treatment of finite Anderson models with up to 15 sites [50].

5.3.3 Fixpoint function

Recalling that the input from the DMFT framework to the impurity solver is the bath Green's function $G_0(i\omega_m)$, see Fig. 5.4, we can now summarize the required steps of the ED algorithm for calculating the interacting Green's function G_I of Eq. (5.2),

1. Determine the finite Anderson model parameters $\{\epsilon_k, V_k\}$ by fitting the corresponding non-interacting Green's function $G_{0,\mathrm{ED}}(i\omega_n)$ to G_0 by minimization of χ^2 in Eq. (5.17)

$$\min_{\epsilon_k, V_k} \chi^2(\epsilon_k, V_k) \rightarrow \{\epsilon_k, V_k\}\,. \tag{5.25}$$

2. Given $\{\epsilon_k, V_k\}$ construct the matrix representation of $\hat{H}_{\mathrm{ED}}$ in the occupation number basis and determine its eigenvectors $|\nu\rangle$ and eigenvalues E_ν by diagonalization

$$\hat{H}_{\mathrm{ED}}|\nu\rangle = E_\nu|\nu\rangle\,. \tag{5.26}$$

3. Use the eigenspectrum of $\hat{H}_{\mathrm{ED}}$ to calculate the interacting impurity Green's function G_I using the Lehman spectral representation Eq. (5.21)

$$G_{I\sigma}(i\omega_n) = \frac{1}{\mathcal{Z}} \sum_{\nu\mu} \frac{|\langle\mu|c_\sigma^\dagger|\nu\rangle|^2}{i\omega_n + E_\nu - E_\mu} \left(e^{-\beta E_\nu} + e^{-\beta E_\mu}\right)\,. \tag{5.27}$$

These steps are schematically shown to the right in Fig. 5.4, replacing the impurity Green's function block in the DMFT coupled equations of Fig. 4.3.

In the view of the fixpoint function formulation of DMFT and Eq. (4.37) in terms of the bath Green's function G_0 the ED algorithm enables a third very interesting parametrization. By breaking the DMFT coupled equations after the χ^2 minimization procedure the resulting fixpoint function depends on the small set of parameters $\mathbf{x} = \{\epsilon_k, V_k\}$

$$\tilde{F}_{U\beta}(G_0) = G_0 \quad \underset{\mathrm{ED}}{\longrightarrow} \quad F_{U\beta}(\mathbf{x}) = \mathbf{x}\,. \tag{5.28}$$

The low dimensional vector-space spanned by $\mathbf{x}$ provides a great simplification when studying the properties of the DMFT fixpoint mapping F, compared to the infinite-dimensional vector-space of Matsubara Green's functions and self-energies ($G_0(i\omega)$ or $\Sigma(i\omega_n)$).

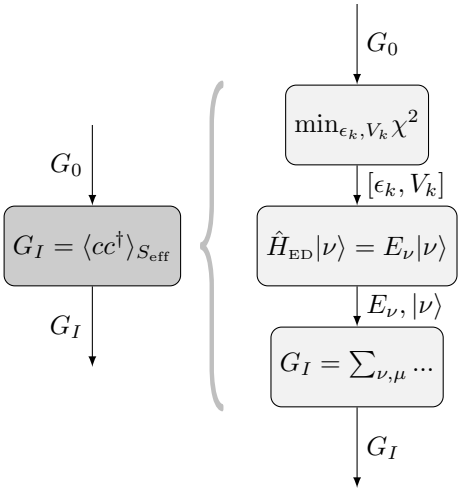


Figure 5.4: Detailed schematic (right) of the Exact Diagonalization impurity solver (left).

In the low dimensional vector-space $\mathbf{x}$ it is straight forward to numerically evaluate the Jacobian J_F of the DMFT fixpoint function $F_{U\beta}(\mathbf{x})$. To this end we have used the modified central finite differences formula

$$J_f(\mathbf{x}) = \nabla_{\mathbf{x}} F_{U\beta}(\mathbf{x}) = \frac{\partial F_{U\beta}(\mathbf{x})}{\partial x_j} \approx \frac{F_{U\beta}(\mathbf{x} + h_j \hat{\mathbf{x}}_j) - F_{U\beta}(\mathbf{x} - h_j \hat{\mathbf{x}}_j)}{2h_j} \,, \tag{5.29}$$

where h_j is the discretization and $\hat{\mathbf{x}}_j$ is the unit-vector in the jth dimension. Due to the spread of the parameters in $\mathbf{x} = \{\epsilon_k, V_k\}$ a relative scaling of h_j was used to stabilize the Jacobian calculation

$$h_j = \bar{h} x_j \,, \tag{5.30}$$

with fixed $\bar{h}$ on the order of 1% to 10%. With access to J_F the DMFT fixpoint problem lends it self to an in-depth analysis using dynamical system methods, providing a deeper understanding of the critical behavior at the Mott metal-insulator transition, as discussed in Paper I.

5.3.4 Implementation

An important part of the work has been the implementation of the Exact Diagonalization routine. The main part of the DMFT-ED calculation was

implemented in the scripting language Python [90]. The benefit of this language is that it comes with "batteries included" in the sense that it comes with an extensive and useful standard library. Especially the numerical modules Numpy and Scipy [91] has simplified and speeded up the development by providing basic array manipulation routines, linear algebra through its LAPACK front-end, multi dimensional root solvers like Broyden's second method (used to solve the DMFT equations) and multidimensional minimizers like the conjugate gradient and the non-linear least squares methods.

As any interpreted language Python comes with a cost, the avoidance of compilation makes explicit loops slow. Numpy solves most of this through its compiled implementation of array operations. But in the case of the speed-critical Lehman summation in Eq. (5.21) the Numpy based implementation did not suffice. After profiling the sum was re-implemented in Fortran and interfaced with Python using the wrapper tool f2py [92].

To further increase the computational throughput the Lehman sum was parallelized over Hilbert space blocks with fixed number of spin up and spin down $(N_\uparrow, N_\downarrow)$. This is possible as the matrix element $\langle\mu|c_\sigma^\dagger|\nu\rangle$ in Eq. (5.21) only connects eigenstates in the block $(N_\sigma, N_{\bar{\sigma}})$ to eigenstates in the block $(N_\sigma{+}1, N_{\bar{\sigma}})$. Using this the summation over eigenstates μ and ν in Eq. (5.21) can be rewritten as

$$\begin{aligned} G_{I\sigma}(i\omega_n) =& \frac{1}{\mathcal{Z}} \sum_{N_\sigma, N_{-\sigma}} \left[\sum_{\substack{\mu\in(N_\sigma, N_{-\sigma}) \\ \nu\in(N_\sigma+1, N_{-\sigma})}} \frac{|\langle\mu|c_\sigma^\dagger|\nu\rangle|^2}{i\omega_n + E_\mu - E_\nu} \left(e^{-\beta E_\nu} + e^{-\beta E_\mu}\right) \right] \\ =& \frac{1}{\mathcal{Z}} \sum_{N_\sigma, N_{-\sigma}} G_{(N_\sigma, N_{-\sigma})}(i\omega_n)\,, \end{aligned} \tag{5.31}$$

where the calculation of the partial Green's function $G_{(N_\sigma, N_{\bar{\sigma}})}(i\omega_n)$ only requires the eigenstates from two blocks. Using the Message Passing Interface (MPI) and the Python module mpi4py the sum over blocks was distributed over MPI ranks. Load balancing was performed through a simple algorithm taking into account the block size of the required blocks for each partial Green's function, the N_Γ^3 scaling of the diagonalization and the remaining Lehman sum.

5.4 Distributional Exact Diagonalization

In Exact Diagonalization the bath Green's function G_0, obtained from the lattice Green's function and self-energy by Eq. (4.4), is approximated by the Green's function of a finite Anderson model $G_{0,\mathrm{ED}}$ defined in Eq. (5.15). The model has only a finite number of poles on the real axis, while the lattice has a continuum of states, so it is not a priori clear how to motivate this approximation. Looking at the spectral functions of G_0 and $G_{0,\mathrm{ED}}$, the ED method attempts to fit a continuous set of states with a finite set of resonances.

However this is a much to harsh view on the method. Exact Diagonalization, applied as an impurity solver in DMFT, should be viewed as the thermodynamical best fit of a finite Anderson model to the continuous bath. This is the case as the actual fitting take place on the finite temperature Matsubara frequencies, not on the real axis. The use of the Matsubara frequency Green's functions turns ED into an approximation optimized for computing thermodynamical averages, like the double occupancy and quasiparticle weight studied in Paper I.

Distributional Exact Diagonalization attempts to address the deficiency of ED in representing the real frequency response functions by considering not just a single Anderson model, but an entire ensemble. A continuous bath Green's function is then readily obtained by taking the ensemble average of many finite models. The basic idea is to use the spectral function A_0 of G_0, given by

$$A_0(\omega) = -\frac{1}{\pi}\mathrm{Im}[G_0(\omega + i0^+)] \tag{5.32}$$

as a probability distribution. An ensemble of bath Green's functions $G_0^{(\nu)}$ containing n poles can then be generated by drawing random pole positions $b_i^{(\nu)}$ having the probability distribution A_0

$$G_0^{(\nu)}(z) = \sum_{i=1}^{n} \frac{a_i^{(\nu)}}{z - b_i^{(\nu)}}\,, \tag{5.33}$$

and requiring that the residues a_i sum to one. The ensemble sample $G_0^{(\nu)}$ can then be mapped directly to a finite Anderson model with $N_\nu - 1$ bath states and solved using ED.

The bath Green's function ensemble average $\bar{G}_0$ of N finite Anderson model realizations

$$\bar{G_0} = \frac{1}{N}\sum_{\nu} G_0^{(\nu)}\,, \tag{5.34}$$

is equal to the original continuous bath Green's function G_0 in the limit

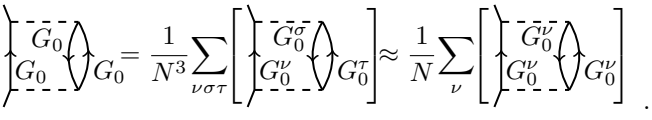
$$\bar{G}_0 \xrightarrow[N\to\infty]{} G_0 \,. \tag{5.35}$$

While the use of ensemble averages trivially reproduce the non-interacting limit its bearing on the interacting system is not directly evident.

In a leap of faith the same ensemble average approximation on the impurity self-energy can be performed

$$\bar{\Sigma}_I = \frac{1}{N}\sum_{\nu} \Sigma_I^{(\nu)} \,. \tag{5.36}$$

This ensemble averaged impurity self-energy is a non-trivial approximation. It corresponds to replacing cross-correlations between different finite Anderson model samples with internal correlations. Consider for example the second order weak coupling diagrammatic expansion of Σ_I that is approximated as



. (5.37)

The ensemble averaging procedure gives this approximation applied to all orders in the self-energy.

It is important to note that the self-energy can not be approximated by applying the Dyson equation to the ensemble averaged interacting and non-interacting Green's functions, i.e. $\bar{\Sigma}_I \not\approx \bar{G}_0^{-1} + \bar{G}^{-1}$, as such a procedure requires the zeros of $\bar{G}_0$ to be zeros of $\bar{G}$ in order to produce a proper self-energy.

5.4.1 Mapping poles to an Anderson model

To map the bath Green's function with a finite number of poles in Eq. (5.33), determined by the coefficients $\{a_i^{(\nu)}, b_i^{(\nu)}\}$, to the an Anderson model with bath greens function

$$G_0^{(\nu)} = \frac{1}{z - \epsilon_0^{(\nu)} - \Delta^{(\nu)}(z)} \,, \tag{5.38}$$

and hybridization function

$$\Delta^{(\nu)}(z) = \sum_k \frac{(V_k^{(\nu)})^2}{z - \epsilon_k^{(\nu)}} \,, \tag{5.39}$$

is a simple algebraic problem. The bath level energies $\epsilon_k^{(\nu)}$ are the roots of G_0

$$G_0^{(\nu)}(\epsilon_k^{(\nu)}) = 0\,, \tag{5.40}$$

determined by the polynomial

$$\sum_i a_i^{(\nu)} \prod_{j\neq i} (z - b_j^{(\nu)}) = 0\,. \tag{5.41}$$

The bath hybridization V_k can be obtained from the derivatives of G_0 at the roots ϵ_k,

$$\partial_z G_0^{(\nu)}(z)|_{z=\epsilon_k^{(\nu)}} = -\frac{1}{(V_k^{(\nu)})^2}\,, \tag{5.42}$$

and the impurity energy level can be obtained from

$$\epsilon_0^{(\nu)} = -(G_0^{(\nu)})^{-1}(z) - \Delta^{(\nu)}(z) + z\,, \tag{5.43}$$

where one can use the fact that G_0 is zero at the poles a_n to get

$$\epsilon_0^{(\nu)} = -\Delta(b_n^{(\nu)}) + b_n^{(\nu)}\,. \tag{5.44}$$

Another possible route for computing ϵ_0 is by comparing the asymptotic $1/z^2$ behaviors of the two forms, such an analysis gives

$$\epsilon_0^{(\nu)} = \sum_j a_i^{(\nu)} b_i^{(\nu)}\,. \tag{5.45}$$

5.4.2 Exploratory study

There are other methods like the Numerical Renormalization Group (NRG) and Density Matrix Renormalization Group (DMRG) solvers that can compute spectral functions on the real axis [93, 94]. These approaches have yet to be generalized to large scale multiorbital models, but the available high quality results for the single band Hubbard model are ideal benchmarks for the Dist-ED approach.

In Paper II we make several comparative studies of Dist-ED finding astonishing good agreement when comparing to NRG and Quantum Monte Carlo calculations. These results should be taken as a strong indication that there is more to be learned from the Dist-ED cross correlation neglecting class of approximations. The exploratory calculations also raised a couple of issues. For weak coupling the replacement of cross correlations with internal

correlations in Eq. (5.37) overestimates the low energy spectral weight. This is caused by an overrepresentation of low energy weight in G_0 that can be compensated for by dropping configurations. A remaining open questions is how to deal with the freedom in choosing the relative residue strengths a_n in the ensemble of Anderson models. They can either be chosen randomly or used to further tune the properties of the finite model. To conclude Dist-ED is an interesting approach and an interesting approximation of the local impurity problem that deserves further attention. By being inherently and trivially parallelizable it holds promise for computing real frequency spectral functions of multi-orbital systems.

Chapter 6

Gutzwiller Variational Method

In the 60s Martin C. Gutzwiller wrote a series of papers introducing a variational wave-function Ψ_G aimed at describing the physics of itinerant fermion systems with local correlation [95–97]. Originally the wave-function Ψ_G was designed exclusively for the single-band Hubbard model, and constructed by combining an uncorrelated Slater determinant wave-function Ψ_0 with a many-body operator $\hat{\mathcal{P}}_G$ that reduced the energetically costly double occupancy locally on every site

$$|\Psi_G\rangle = \hat{\mathcal{P}}_G|\Psi_0\rangle\,. \tag{6.1}$$

However evaluating expectation values even for this seemingly simple wave-function is a non-trivial many-body problem. Guided by intuition Gutzwiller therefore suggested an approximation enabling analytic evaluation of expectation values [97].

It was not until two decades later, in 1989, that Metzner and Vollhardt [29] realized that fermionic systems have a non-trivial limit in infinite dimensions. With this insight they could show that the proposal by Gutzwiller was actually exact in the limit of infinite dimensions. These ideas started an entire sub-field in the strong-correlations community. Today Gutzwiller's proposed wave-function and its generalizations bear his name and are simply called *Gutzwiller wave-functions*, and the approximation scheme is referred to as the *Gutzwiller Approximation*.

The generalization of the Gutzwiller wave-function and the Gutzwiller approximation to more realistic multi-band Hubbard models, with general

local interactions, turned out to be involved. The issue was solved in 1998 by Bünemann *et al.* [98], whose initial theoretical paper was later followed by a throughout calculation for the correlated transition metal Ni [99–101], and more recently a study of antiferromagnetism in the iron pnictides [102].

However, method papers on numerical methods for solving the multi-band Gutzwiller problem has been scarce. A powerful approach has been developed by Deng *et al.* [103, 104], albeit limited to density-density interactions. In Paper III we generalize Deng's ideas to general interactions. After publishing Paper III on the ArXiv, another paper by Bünemann *et al.* appeared with details on their numerical procedure [105]. Although both methods solve the multi-orbital Gutzwiller problem the two numerical approaches differ. The idea of Bünemann *et al.* [105] is that of a direct minimization while our method in Paper III directly exploits the separability of the problem.

In 2007 Lechermann *et al.* [106] published a generalization of the Kotliar-Ruckenstein slave-boson formalism [107] to multi-band Hubbard models. The final equations was later shown to be identically equal to the equations of the corresponding Gutzwiller problem [108]. However, even though the slave-boson approach in the end gives the same variational problem as the Gutzwiller method its derivation takes a very different path, and opens up for different interpretations and extensions.

One such important line of thought is the Gutzwiller formalism developed by Fabrizio and Lanatà [109–111], who realized the merit of rewriting the Gutzwiller problem in a slave-boson like form. This was done by replacing the projector $\hat{\mathcal{P}}_G$ variational variable with an associated operator ϕ that encodes all local properties of Ψ_G. One of the benefits of this rewrite is that local expectation values no longer explicitly depends on the Slater determinant. This is one important step in separating the problem which is one of the main topics of Papers III and V.

This chapter has the following structure; Section 6.1 treats the general Gutzwiller wave-function together with the corresponding variational minimization problem that can be tackled analytically using the Gutzwiller approximation described in Section 6.2. By the introduction of the mixed basis representation in Section 6.3 the Gutzwiller minimization problem is partially separated in local and non-local parts. How wave-function symmetries translate to this form is derived in Section 6.4 and it also enables direct calculation of the entanglement entropy as outlined in Section 6.5. We then go on to formulate our minimization algorithm by introducing a vector-space representation in Section 6.6 that is used to completely separate the local and non-local minimization problems in Section 6.7. The

insights from the separation of the problem are then used in Section 6.8 to formulate an efficient Lagrange based solver.

6.1 Gutzwiller wave-function

We will adopt the multi-orbital formulation of the Gutzwiller wave-function using the notation of Fabrizio and Lanatà [109–111]. The Gutzwiller wave-function takes the general form of Eq. (6.1) where Ψ_0 is assumed to be an uncorrelated Slater determinant wave-function

$$|\Psi_0\rangle = \prod_{\epsilon_{\mathbf{k}n}<\epsilon_F} c^\dagger_{\mathbf{k}n}|0\rangle\,, \tag{6.2}$$

filled up to the Fermi surface of some non-interacting Hamiltonian with dispersion $\epsilon_{\mathbf{k}n}$ and Fermi level ϵ_F. The operator $\hat{\mathcal{P}}_G$ is constructed as a product of local operators $\hat{\mathcal{P}}_{\mathbf{R}}$ acting separately on each lattice site

$$\hat{\mathcal{P}}_G = \prod_{\mathbf{R}} \hat{\mathcal{P}}_{\mathbf{R}}\,, \tag{6.3}$$

where $\hat{\mathcal{P}}_{\mathbf{R}}$ in turn is a completely general operator in the local Fock-space that can be written, without any loss of generality, as

$$\hat{\mathcal{P}}_{\mathbf{R}} = \sum_{\Gamma\Gamma'} \lambda_{\Gamma\Gamma'}(\mathbf{R})|\Gamma,\mathbf{R}\rangle\langle\Gamma',\mathbf{R}|\,, \tag{6.4}$$

where $|\Gamma,\mathbf{R}\rangle$ are the local many-body occupation number states, and $\lambda_{\Gamma\Gamma'}(\mathbf{R})$ are our variational parameters, that we will refer to as *configuration weights*. The local many-body occupation number states $|\Gamma,\mathbf{R}\rangle$ can in turn be constructed as

$$|\Gamma,\mathbf{R}\rangle = \left(\prod_{\alpha\in\Gamma} c^\dagger_{\mathbf{R}\alpha}\right)|0,\mathbf{R}\rangle\,, \tag{6.5}$$

using the site local vacuum state $|0,\mathbf{R}\rangle$, and the fermion creation operator $c^\dagger_{\mathbf{R}\alpha}$, creating an electron on site $\mathbf{R}$ with composite quantum-number α (including all local spin and orbital degrees of freedom). In the following treatment we are going to refer to $\hat{\mathcal{P}}_{\mathbf{R}}$ as the *Gutzwiller projector*.[1] To conclude, the Gutzwiller wave-function Ψ_G is parametrized by Ψ_0 and λ, $\Psi_G = \Psi_G(\Psi_0,\lambda)$.

[1] Note however that $\hat{\mathcal{P}}_{\mathbf{R}}$ is not a projection operator in the mathematical sense, as $\hat{\mathcal{P}}_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}} \neq \hat{\mathcal{P}}_{\mathbf{R}}$.

Our interest lies in using Ψ_G and the variational principle [112], to obtain an approximate ground state for multi-orbital Hubbard models $\hat{H}$ with the general form

$$\hat{H} = \hat{\mathcal{T}} + \hat{H}_{int} = \frac{1}{\mathcal{N}} \sum_{\mathbf{R}\mathbf{R}'} \sum_{\alpha\beta} t^{\alpha\beta}_{\mathbf{R}\mathbf{R}'} c^{\dagger}_{\mathbf{R}\alpha} c_{\mathbf{R}'\beta} + \frac{1}{\mathcal{N}} \sum_{\mathbf{R}} \hat{H}_{loc}(\mathbf{R}) \,, \tag{6.6}$$

with a kinetic term $\hat{\mathcal{T}}$ containing single-particle hopping $t^{\alpha\beta}_{\mathbf{R}\mathbf{R}'}$ from states with composite quantum numbers β at site $\mathbf{R}'$ to states α at $\mathbf{R}$, and an interaction $\hat{H}_{int}$ term composed by purely local many-body interactions $\hat{H}_{loc}(\mathbf{R})$, and $\mathcal{N}$ is the number of lattice sites.[2]

The variational principle gives an upper bound to the ground state energy E_0 in terms of the minimization program

$$\begin{aligned} E_0 \leq \min_{\Psi_G} \{ \langle \Psi_G | \hat{H} | \Psi_G \rangle : \langle \Psi_G | \Psi_G \rangle = 1 \} \\ \equiv \min_{\Psi_0, \lambda} \{ \langle \Psi_0 | \hat{\mathcal{P}}^{\dagger} \hat{H} \hat{\mathcal{P}} \ | \Psi_0 \rangle : \langle \Psi_0 | \hat{\mathcal{P}}^{\dagger} \hat{\mathcal{P}} \ | \Psi_0 \rangle = 1 \} \,. \end{aligned} \tag{6.7}$$

Thus, we seek to minimize the expectation value $\langle \Psi_G | \hat{H} | \Psi_G \rangle$ with respect to Ψ_0 and λ, under the constraint that Ψ_G is normalized $\langle \Psi_G | \Psi_G \rangle = 1$. Unfortunately, exact evaluation of $\langle \Psi_G | \hat{H} | \Psi_G \rangle$ is in itself a many-body problem requiring Wick contractions [113] to infinite orders, which can in principle be performed using variational Monte Carlo [114, 115].

6.2 Gutzwiller approximation

To circumvent the many-body problem posed by the expectation value of Ψ_G in Eq. (6.7), Gutzwiller proposed an approximation scheme enabling analytic calculations [97], without explicitly using the limit of infinite dimensions. Gutzwiller did actually realize that, even though his analytic approximation failed in predicting a ferromagnetic ground state in one dimension, its relevance would improve in two and three dimensions [97]. Metzner and Vollhardt where later able to prove that Gutzwiller's approximation becomes exact in the infinite dimensional limit [29].

In our generalized multi-band formalism this Gutzwiller Approximation

[2] The slightly unorthodox $1/\mathcal{N}$ scaling of $\hat{H}$ simplifies the notation by making subsequent kinetic expectation values $\langle \hat{\mathcal{T}} \rangle$ intensive rather than extensive, see Section 2.1.

takes the form of the constraints [109]

$$\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle = 1\,, \tag{6.8}$$

$$\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}c^\dagger_{\alpha\mathbf{R}}c_{\beta\mathbf{R}}|\Psi_0\rangle = \langle\Psi_0|c^\dagger_{\alpha\mathbf{R}}c_{\beta\mathbf{R}}|\Psi_0\rangle\,, \tag{6.9}$$

in combination with the limit of infinite dimensions. We will henceforth refer to Eqs. (6.8) and (6.9) as GA1 and GA2 respectively. Under these assumptions analytic expressions for $\langle\Psi_G|\hat{H}|\Psi_G\rangle$ can readily be derived. As shown in Appendix B, imposing GA1 and GA2, and taking the limit of infinite dimensions, results in a tremendous cancellation of Gutzwiller projector contributions to expectation values.

For a general local many-body operator $\hat{\mathcal{O}}(\mathbf{R})$ acting on site $\mathbf{R}$, all contractions of Gutzwiller projector pairs $\hat{\mathcal{P}}^\dagger_{\mathbf{R}'}\hat{\mathcal{P}}_{\mathbf{R}'}$ acting on other sites $\mathbf{R}' \neq \mathbf{R}$ are zero [Eq. (B.34)] and results in

$$\begin{aligned}&\langle\Psi_G|\hat{\mathcal{O}}(\mathbf{R})|\Psi_G\rangle\\ &\quad= \langle\Psi_0|\Big(\prod_{\mathbf{R}'}\hat{\mathcal{P}}^\dagger_{\mathbf{R}'}\Big)\hat{\mathcal{O}}(\mathbf{R})\Big(\prod_{\mathbf{R}''}\hat{\mathcal{P}}_{\mathbf{R}''}\Big)|\Psi_0\rangle = \langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle\,.\end{aligned} \tag{6.10}$$

The same results holds for quadratic operators acting on two different sites $\mathbf{R}$ and $\mathbf{R}'$ [Eq. (B.35)] and the remaining operators can be ordered as

$$\langle\Psi_G|c^\dagger_{\mathbf{R}\alpha}c_{\mathbf{R}'\beta}|\Psi_G\rangle = \langle\Psi_0|[\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{P}}_{\mathbf{R}}][\hat{\mathcal{P}}^\dagger_{\mathbf{R}'}c_{\mathbf{R}'\beta}\hat{\mathcal{P}}_{\mathbf{R}'}]|\Psi_0\rangle\,. \tag{6.11}$$

The remaining non-trivial contractions in Eq. (6.11) can be absorbed in a renormalization matrix $\mathcal{R}$, see Eq. (B.40), giving non-local quadratic expectation values on the form

$$\begin{aligned}&\langle\Psi_0|[\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{P}}_{\mathbf{R}}][\hat{\mathcal{P}}^\dagger_{\mathbf{R}'}c_{\mathbf{R}'\beta}\hat{\mathcal{P}}_{\mathbf{R}'}]|\Psi_0\rangle\\ &\qquad\qquad= \sum_{\gamma\delta}\mathcal{R}_{\gamma\alpha}(\mathbf{R})\langle\Psi_0|c^\dagger_{\mathbf{R}\gamma}c_{\mathbf{R}'\delta}|\Psi_0\rangle\mathcal{R}^\dagger_{\beta\delta}(\mathbf{R}')\,,\end{aligned} \tag{6.12}$$

where $\mathcal{R}$ can be obtained [Eq. (B.44)] from purely local expectation values through

$$\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\beta}\hat{\mathcal{P}}_{\mathbf{R}}c_{\mathbf{R}\alpha}|\Psi_0\rangle = \sum_{\gamma}\mathcal{R}_{\gamma\beta}(\mathbf{R})\langle\Psi_0|c^\dagger_{\mathbf{R}\gamma}c_{\mathbf{R}\alpha}|\Psi_0\rangle\,. \tag{6.13}$$

Now, as every contribution to the kinetic energy $\hat{\mathcal{T}}$ in Eq. (6.6) will be affected by $\mathcal{R}$, it is convenient to define the renormalized kinetic operator

[Eq. (B.42)]

$$\hat{\mathcal{T}}^G(\mathcal{R}) = \frac{1}{\mathcal{N}} \sum_{\mathbf{R}\mathbf{R}'} \sum_{\alpha\beta} \tilde{t}^{\alpha\beta}_{\mathbf{R}\mathbf{R}'}(\mathcal{R}) c^\dagger_{\mathbf{R}\alpha} c_{\mathbf{R}'\beta} \,, \tag{6.14}$$

where $\mathcal{R}$ is absorbed in the renormalized hopping [Eq. (B.43)]

$$\tilde{t}^{\alpha\beta}_{\mathbf{R}\mathbf{R}'} = \sum_{\gamma\delta} \mathcal{R}_{\alpha\gamma}(\mathbf{R}) t^{\gamma\delta}_{\mathbf{R}\mathbf{R}'} \mathcal{R}^\dagger_{\delta\beta}(\mathbf{R}') \,. \tag{6.15}$$

Thus applying the Gutzwiller approximation in terms of Eqs. (6.10) – (6.15) on the total energy expectation value $\langle \Psi_G | \hat{H} | \Psi_G \rangle$ gives the energy expression

$$\mathcal{E} = \langle \Psi_0 | \hat{\mathcal{T}}^G(\mathcal{R}) | \Psi_0 \rangle + \langle \Psi_0 | \hat{\mathcal{P}}^\dagger_{\mathbf{R}} \hat{H}_{\mathrm{loc}}(\mathbf{R}) \hat{\mathcal{P}}_{\mathbf{R}} | \Psi_0 \rangle \,, \tag{6.16}$$

where we have dropped the sum over all sites in the local interaction energy by assuming translational invariance. The variational problem of Eq. (6.7) can then, in the Gutzwiller approximation, be stated as the minimization of $\mathcal{E}$ in Eq. (6.16)

$$\begin{aligned} \min_{\lambda,\Psi_0} \{ \mathcal{E}[\lambda, \Psi_0] : \langle \Psi_0 | \Psi_0 \rangle = 1, \, \langle \Psi_0 | \hat{\mathcal{P}}^\dagger_{\mathbf{R}} \hat{\mathcal{P}}_{\mathbf{R}} | \Psi_0 \rangle = 1, \\ \langle \Psi_0 | \hat{\mathcal{P}}^\dagger_{\mathbf{R}} \hat{\mathcal{P}}_{\mathbf{R}} c^\dagger_{\alpha\mathbf{R}} c_{\beta\mathbf{R}} | \Psi_0 \rangle = \langle \Psi_0 | c^\dagger_{\alpha\mathbf{R}} c_{\beta\mathbf{R}} | \Psi_0 \rangle \} \,, \end{aligned} \tag{6.17}$$

under the constraints GA1, GA2, and Ψ_0 normalization. In this formulation the mathematically most involved inter-dependency is hidden in the renormalization matrix $\mathcal{R}$, that enters into the kinetic energy through Eqs. (6.14) and (6.15). The renormalization matrix $\mathcal{R}$ in itself is a function of both Ψ_0 and $\hat{\mathcal{P}}_{\mathbf{R}}$ by virtue of Eq. (6.13), $\mathcal{R} = \mathcal{R}(\Psi_0, \hat{\mathcal{P}}_{\mathbf{R}})$.

In the following sections we will systematically rewrite and reduce the variational problem of Eq. (6.17). For multi-orbital models there is much to gain from a formulation that lends itself to an efficient numerical implementation.

6.3 Mixed basis representation

Our first step in reducing the variational problem is to pass over to the *mixed basis representation* introduced by Lanatà *et al.* [110]. Through this change of variables in the variational problem we are going to decouple all local expectation values from the Slater determinant Ψ_0.

We begin by exploiting that the variational problem is invariant under local unitary transformations of the single-particle operators $c^\dagger_{\mathbf{R}\alpha}$. This freedom can be gauged away by introducing a set of natural basis operators $d^\dagger_{\mathbf{R}\alpha}$ for the Slater determinant Ψ_0. The operators $d^\dagger_{\mathbf{R}\alpha}$ are uniquely defined (up to a phase) as the unitary transform $\mathcal{U}$ of $c^\dagger_{\mathbf{R}\alpha}$

$$d^\dagger_{\mathbf{R}\alpha} = \sum_\beta \mathcal{U}_{\alpha\beta} c^\dagger_{\mathbf{R}\beta} \,, \tag{6.18}$$

that diagonalize the local single-particle density matrix of the Slater determinant Ψ_0

$$\langle \Psi_0 | d^\dagger_{\mathbf{R}\alpha} d_{\mathbf{R}\beta} | \Psi_0 \rangle = \delta_{\alpha\beta} n^0_\alpha \,, \tag{6.19}$$

whose diagonal components will be denoted by n^0_α. Expressing the renormalization matrix $\mathcal{R}$ in this basis cancels the sum in its original defining expression Eq. (6.13) and gives the simplified relation

$$\langle \Psi_0 | \hat{\mathcal{P}}^\dagger_{\mathbf{R}} c^\dagger_{\mathbf{R}\beta} \hat{\mathcal{P}}_{\mathbf{R}} d_{\mathbf{R}\alpha} | \Psi_0 \rangle = \mathcal{R}_{\alpha\beta}(\mathbf{R}) \langle \Psi_0 | d^\dagger_{\mathbf{R}\alpha} d_{\mathbf{R}\alpha} | \Psi_0 \rangle \,. \tag{6.20}$$

Along the same line it is also possible to rewrite the constraints in the Gutzwiller variational problem Eq. (6.17) in terms of $d_{\mathbf{R}\alpha}$.

To proceed with the derivation of the mixed basis representation we are going to take a slightly different route from the original derivation [110], and use the exponential form for the local many-body density matrix $\hat{\rho}_0$ of Ψ_0 to arrive at the necessary equations. Formally, the local many-body density matrix $\hat{\rho}_0$ at site $\mathbf{R}$ is defined as the partial trace of the full many-body density matrix $|\Psi_0\rangle\langle\Psi_0|$ over all other sites,[3]

$$\hat{\rho}_0(\mathbf{R}) = \mathrm{Tr}_{\mathbf{R}' \neq \mathbf{R}} [|\Psi_0\rangle\langle\Psi_0|] \,. \tag{6.21}$$

However, the reduced many-body density matrix $\hat{\rho}_0$ of a Slater determinant is as a matter of fact completely determined by its corresponding single-particle density matrix. In the natural basis, where the local single-particle density matrix is diagonal [Eq. (6.19)], $\hat{\rho}_0$ is directly given by the exponential operator [116]

$$\hat{\rho}_0(\mathbf{R}) = \frac{1}{\mathcal{Z}} e^{-\sum_\alpha \epsilon_\alpha d^\dagger_{\mathbf{R}\alpha} d_{\mathbf{R}\alpha}} \,, \tag{6.22}$$

where the normalization factor $\mathcal{Z}$ and ϵ_α are determined by the local density matrix $n^0_\alpha(\mathbf{R})$

$$\mathcal{Z}^{-1} = \prod_\alpha (1 - n^0_\alpha) \,, \quad \epsilon_\alpha = \ln \left(\frac{1 - n^0_\alpha}{n^0_\alpha} \right) \,. \tag{6.23}$$

[3] I.e. tracing out the many-body states of the rest of the system in a proper Schmidt decomposition of $|\Psi_0\rangle\langle\Psi_0|$ [116].

This definition shows that $\hat{\rho}_0$ is also diagonal, when expressed in the many-body occupation number basis $|\eta\rangle$ of the natural basis operators $d_\alpha^\dagger$. Furthermore all density matrices are by definition positive semi-definite. These two properties of $\hat{\rho}_0$ make it possible to directly perform a Cholesky factorization

$$\hat{\rho}_0(\mathbf{R}) = \phi_0 \phi_0^\dagger\,, \text{ where } \phi_0 = \frac{1}{\sqrt{\mathcal{Z}}} e^{-\frac{1}{2}\sum_\alpha \epsilon_\alpha d_{\mathbf{R}\alpha}^\dagger d_{\mathbf{R}\alpha}}\,. \tag{6.24}$$

The exponential form of Eq. (6.24) also makes it possible to derive the commutation relations between ϕ_0 and $d_\alpha^\dagger$, d_α

$$d_{\mathbf{R}\alpha}\phi_0 = e^{-\frac{\epsilon_\alpha}{2}}\phi_0\, d_{\mathbf{R}\alpha} = \sqrt{\frac{n_\alpha^0}{1-n_\alpha^0}}\,\phi_0\, d_{\mathbf{R}\alpha}, \quad d_{\mathbf{R}\alpha}^\dagger\phi_0 = \sqrt{\frac{1-n_\alpha^0}{n_\alpha^0}}\,\phi_0\, d_{\mathbf{R}\alpha}^\dagger\,, \tag{6.25}$$

where $\phi_0^\dagger = \phi_0$ was used to obtain the last relation.

It is of course also possible to write ϕ_0 as

$$\phi_0 = \sum_\eta \phi_{0,\eta} |\eta, \mathbf{R}\rangle\langle\eta, \mathbf{R}|\,, \tag{6.26}$$

using projectors in the many-body occupation number basis $|\eta, \mathbf{R}\rangle$ of the natural basis

$$|\eta, \mathbf{R}\rangle = \prod_{\alpha\in\eta} d_{\mathbf{R}\alpha}^\dagger |0, \mathbf{R}\rangle\,. \tag{6.27}$$

The coefficients $\phi_{0,\eta}$ do of course only depend on n_α^0 and are given by

$$\phi_{0,\eta} = \prod_\alpha (\sqrt{1-n_\alpha^0})^{1-n_\alpha(\eta)} (\sqrt{n_\alpha^0})^{n_\alpha(\eta)}\,, \tag{6.28}$$

where $n_\alpha(\eta) = \langle\eta| d_\alpha^\dagger d_\alpha |\eta\rangle$.

The question that now arises is of course whether we also can write down the Cholesky factorization of the local many-body density matrix $\hat{\rho}$ of the Gutzwiller wave-function Ψ_G. This proves to actually be quite straight forward. Equipped with ϕ_0 and the Gutzwiller Approximation we obtain

$$\begin{aligned}\hat{\rho}(\mathbf{R}) = \mathrm{Tr}_{\mathbf{R}'\neq\mathbf{R}}[|\Psi_G\rangle\langle\Psi_G|] = \\ \hat{\mathcal{P}}_{\mathbf{R}} \mathrm{Tr}_{\mathbf{R}'\neq\mathbf{R}}[|\Psi_0\rangle\langle\Psi_0|]\hat{\mathcal{P}}_{\mathbf{R}}^\dagger = \hat{\mathcal{P}}_{\mathbf{R}}\phi_0\phi_0^\dagger\hat{\mathcal{P}}_{\mathbf{R}}^\dagger = \phi\phi^\dagger\,,\end{aligned} \tag{6.29}$$

where we have identified $\phi = \hat{\mathcal{P}}_{\mathbf{R}}\phi_0$ as the corresponding Cholesky factorization of the local many-body Gutzwiller density matrix $\hat{\rho} = \phi\phi^\dagger$.

Let us now investigate the object ϕ in detail. As stressed in [110] the operator ϕ is expressed in a "mixed basis", in the sense that it is a combination of $\hat{\mathcal{P}}_{\mathbf{R}}$ written in terms of $c^{\dagger}_{\mathbf{R}\alpha}$ [Eqs. (6.4) and (6.5)] and ϕ_0 expressed in $d^{\dagger}_{\mathbf{R}\alpha}$ [Eq. (6.24)]. This is even more clear if we use the explicit expression in Eq. (6.4) for $\hat{\mathcal{P}}_{\mathbf{R}}$ in terms of the configuration weights λ and Eq. (6.26) for ϕ_0. We then obtain ϕ as

$$\phi \equiv \hat{\mathcal{P}}_{\mathbf{R}}\phi_0 = \sum_{\Gamma\eta} \phi_{\Gamma\eta} |\Gamma, \mathbf{R}\rangle\langle\eta, \mathbf{R}| \,, \tag{6.30}$$

where $\phi_{\Gamma\eta} = \sum_{\Gamma'} \lambda_{\Gamma\Gamma'}\phi_{0,\eta}\langle\Gamma', \mathbf{R}|\eta, \mathbf{R}\rangle$.

The beauty of writing it like this is that, as we are at the liberty of choosing our variational parameters, we can now decide to work with ϕ instead of λ. By making this choice we eliminate the gauge freedom of unitary transforms of the Slater determinant single-particle density matrix, and on top of that ϕ fully determines the local density matrix $\hat{\rho}$ of Ψ_G.

These merits can better be seen by rewriting the variational program of Eq. (6.17) in terms of ϕ and Ψ_0. To accomplish this we need to write all expectation values containing $\hat{\mathcal{P}}_{\mathbf{R}}$ in Eq. (6.17) as traces of ϕ. The following derivation follows the ideas of Lanatà [110, 117], but extends its validity to the limits $n^0_\alpha = 0$, and 1, by using the operator form of ϕ_0 in Eq. (6.24).

Starting with the left-hand side of the GA1 normalization constraint of Eq. (6.8) we get

$$\langle\Psi_0|\hat{\mathcal{P}}^{\dagger}_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle = \mathrm{Tr}[\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle\langle\Psi_0|\hat{\mathcal{P}}^{\dagger}_{\mathbf{R}}] = \mathrm{Tr}[\phi\phi^{\dagger}] \,, \tag{6.31}$$

by using the trace invariance under cyclic permutations. In the natural basis the left-hand side of the GA2 density constraint [Eq. (6.9)] takes the form

$$\begin{aligned}
&\langle\Psi_0|\hat{\mathcal{P}}^{\dagger}_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}|\Psi_0\rangle \\
&\qquad = \mathrm{Tr}[\phi_0\phi_0^{\dagger}\hat{\mathcal{P}}^{\dagger}_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}] = \frac{\sqrt{n^0_\beta(1-n^0_\alpha)}}{\sqrt{n^0_\alpha(1-n^0_\beta)}}\mathrm{Tr}[\phi^{\dagger}\phi d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}] \,,
\end{aligned} \tag{6.32}$$

where we have used the commutation relations in Eq. (6.25) in the last step. Now performing the same steps in the right-hand side of Eq. (6.9) gives

$$\langle\Psi_0|d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}|\Psi_0\rangle = \mathrm{Tr}[\phi_0\phi_0^{\dagger}d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}] = \frac{\sqrt{n^0_\beta(1-n^0_\alpha)}}{\sqrt{n^0_\alpha(1-n^0_\beta)}}\mathrm{Tr}[\phi_0^{\dagger}\phi_0 d^{\dagger}_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}] \,. \tag{6.33}$$

Combining the left and right hand sides [Eqs. (6.32) and (6.33)] the square-root prefactors cancels, giving the density constraint as

$$\mathrm{Tr}[\phi^\dagger \phi d^\dagger_{\alpha\mathbf{R}} d_{\beta\mathbf{R}}] = \mathrm{Tr}[\phi_0^\dagger \phi_0 d^\dagger_{\alpha\mathbf{R}} d_{\beta\mathbf{R}}]\,. \tag{6.34}$$

Note that since ϕ_0 is diagonal and real we have $\phi_0 = \phi_0^\dagger$ and can without ambiguity still use the Slater determinant density matrix on the right-hand side

$$\mathrm{Tr}[\phi_0^\dagger \phi_0 d^\dagger_{\alpha\mathbf{R}} d_{\beta\mathbf{R}}] = \mathrm{Tr}[\phi_0 \phi_0^\dagger d^\dagger_{\alpha\mathbf{R}} d_{\beta\mathbf{R}}] = \langle \Psi_0 | d^\dagger_{\alpha\mathbf{R}} d_{\beta\mathbf{R}} | \Psi_0 \rangle\,. \tag{6.35}$$

We must also rewrite the expression for the renormalization matrix $\mathcal{R}$ in Eq. (6.13). The left-hand side written in terms of ϕ and the natural basis becomes

$$\begin{aligned}\langle \Psi_0 | \hat{\mathcal{P}}^\dagger_{\mathbf{R}} c^\dagger_{\mathbf{R}\beta} \hat{\mathcal{P}}_{\mathbf{R}} d_{\mathbf{R}\alpha} | \Psi_0 \rangle \\ = \mathrm{Tr}[\phi_0 \phi_0^\dagger \hat{\mathcal{P}}^\dagger_{\mathbf{R}} c^\dagger_{\mathbf{R}\beta} \hat{\mathcal{P}}_{\mathbf{R}} d_{\mathbf{R}\alpha}] = \sqrt{\frac{n^0_\alpha}{1-n^0_\alpha}}\,\mathrm{Tr}[\phi^\dagger c^\dagger_{\mathbf{R}\beta} \phi d_{\mathbf{R}\alpha}]\,,\end{aligned} \tag{6.36}$$

and the left-hand side is simplified by the diagonal natural basis single-particle density matrix [Eqs. (6.19), (6.13) and (6.20)]

$$\sum_\gamma \mathcal{R}_{\gamma\beta} \langle \Psi_0 | d^\dagger_{\gamma\mathbf{R}} d_{\alpha\mathbf{R}} | \Psi_0 \rangle = \mathcal{R}_{\alpha\beta}\, n^0_\alpha\,. \tag{6.37}$$

Combining left and right-hand sides gives

$$\mathrm{Tr}[\phi^\dagger c^\dagger_{\mathbf{R}\beta} \phi d_{\mathbf{R}\alpha}] = \sqrt{n^0_\alpha (1-n^0_\alpha)}\,\mathcal{R}_{\alpha\beta}\,, \tag{6.38}$$

a much simplified expression compared to the original Eq. (6.13). Interestingly, the equation for the renormalization-matrix $\mathcal{R}$ can be rewritten even more symmetrically in terms of ϕ_0 by noting that

$$n^0_\alpha = \mathrm{Tr}[\phi_0 \phi_0^\dagger d^\dagger_{\mathbf{R}\alpha} d_{\mathbf{R}\alpha}] = \sqrt{\frac{n^0_\alpha}{1-n^0_\alpha}}\,\mathrm{Tr}[\phi_0^\dagger d^\dagger_{\mathbf{R}\alpha} \phi_0 d_{\mathbf{R}\alpha}]\,. \tag{6.39}$$

Thus the renormalization matrix $\mathcal{R}$ is the ratio of the traces

$$\mathrm{Tr}[\phi^\dagger c^\dagger_{\mathbf{R}\beta} \phi d_{\mathbf{R}\alpha}] = \mathcal{R}_{\alpha\beta} \mathrm{Tr}[\phi_0^\dagger d^\dagger_{\mathbf{R}\alpha} \phi_0 d_{\mathbf{R}\alpha}]\,, \tag{6.40}$$

only differing in the operator ($c^\dagger$ or $d^\dagger$) that is sandwiched between ϕ and ϕ_0.

The last term requiring our attention is the expectation value of the local interaction Hamiltonian $\hat{H}_{loc}(\mathbf{R})$ in Eq. (6.16). As any expectation value of a local operator is the trace of the product of the operator and the local density matrix $\hat{\rho}$ we get

$$\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{H}_{\mathrm{loc}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle = \mathrm{Tr}[\hat{H}_{\mathrm{loc}}\hat{\rho}(\mathbf{R})] = \mathrm{Tr}[\hat{H}_{\mathrm{loc}}\phi\phi^\dagger]\,. \tag{6.41}$$

By now we have expressed all parts of the original variational problem [Eq. (6.17)] in terms of Ψ_0 and ϕ, and can write down the corresponding minimization problem

$$\begin{aligned}\min_{\phi,\Psi_0}\{\mathcal{E}[\phi,\Psi_0] = \langle\Psi_0|\hat{\mathcal{T}}^G[\mathcal{R}]|\Psi_0\rangle + \mathrm{Tr}[\hat{H}_{\mathrm{loc}}\phi\phi^\dagger] :\\ \langle\Psi_0|\Psi_0\rangle = 1,\, \langle\Psi_0|d^\dagger_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}|\Psi_0\rangle = 0, \alpha\neq\beta,\\ \mathrm{Tr}[\phi\phi^\dagger] = 1,\, \mathrm{Tr}[\phi^\dagger\phi d^\dagger_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}] = \langle\Psi_0|d^\dagger_{\alpha\mathbf{R}}d_{\beta\mathbf{R}}|\Psi_0\rangle\}\,,\end{aligned} \tag{6.42}$$

where $\mathrm{Tr}[\phi^\dagger c^\dagger_{\mathbf{R}\beta}\phi d_{\mathbf{R}\alpha}] = \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta}$ and $\langle\Psi_0|d^\dagger_{\alpha\mathbf{R}}d_{\alpha\mathbf{R}}|\Psi_0\rangle = n^0_\alpha$.

So what are the merits of the mixed basis variational problem formulation in Eq. (6.42)? The diagonal constraint on the Ψ_0 single-particle density matrix causes ϕ to automatically absorb the freedom of unitary transforms of fermion operators, and the expression for the renormalization-matrix $\mathcal{R}$ [Eq. (6.38)], is now a single equation instead of a matrix problem. Another seemingly trivial point, that will become very important, is that the interaction energy contribution to the total energy $\mathcal{E}$ no longer depends on Ψ_0, due to Eq. (6.41). This is the first step towards obtaining a fully separated problem where Ψ_0 and ϕ can be optimized independently. The final step is the subject of Section 6.6.

6.4 Wave-function symmetries

In terms of the new parametrization of the Gutzwiller wave-function Ψ_G using the Slater determinant Ψ_0 and the local many-body density matrix factorization $\hat{\rho} = \phi\phi^\dagger$ it is not directly evident how wave-function symmetries apply to ϕ.

Let $\hat{\mathcal{O}}(\mathbf{R})$ be a local operator under which the Gutzwiller wave-function is invariant $\hat{\mathcal{O}}(\mathbf{R})|\Psi_G\rangle = |\Psi_G\rangle$. This is true if the local projector $\hat{\mathcal{P}}_{\mathbf{R}}$ commutes with $\hat{\mathcal{O}}(\mathbf{R})$

$$[\hat{\mathcal{P}}_{\mathbf{R}}, \hat{\mathcal{O}}(\mathbf{R})] = 0\,, \tag{6.43}$$

and if the Slater determinant wave-function also is invariant with respect to $\hat{\mathcal{O}}(\mathbf{R})$, i.e. $\hat{\mathcal{O}}(\mathbf{R})|\Psi_0\rangle = |\Psi_0\rangle$. Locally this is equivalent to that the local many-body density matrix $\hat{\rho}_0$ of Ψ_0 commutes with $\hat{\mathcal{O}}(\mathbf{R})$

$$[\hat{\rho}_0, \hat{\mathcal{O}}(\mathbf{R})] = 0\,. \tag{6.44}$$

The commutation relation directly implies that any integer power α of $\hat{\rho}_0$ fulfill the same commutation relation. Non-integer powers $\alpha > 0$ can be directly defined in terms of their Taylor expansions in integer powers. Thus, also the square root of $\hat{\rho}_0 = \phi_0\phi_0$ commute with the operator

$$[\phi_0, \hat{\mathcal{O}}(\mathbf{R})] = 0\,. \tag{6.45}$$

Now as the factorization of $\hat{\rho} = \phi\phi^\dagger$ is by Eq. (6.29) given as $\phi = \hat{\mathcal{P}}_{\mathbf{R}}\phi_0$, also ϕ commutes with the invariant operator

$$[\phi, \hat{\mathcal{O}}(\mathbf{R})] = 0\,. \tag{6.46}$$

In the case of finite symmetry groups, like the lattice point group $\mathcal{G}$ its transform law of the single particle states

$$g^{-1}\psi_{\mathbf{R}m\sigma}(\mathbf{r})g = \sum_{m'} \mathcal{D}^{(l)}_{mm'}(g)\psi_{\mathbf{R}m'\sigma}(\mathbf{r})\,, \tag{6.47}$$

with $g \in \mathcal{G}$, defines an unitary operator representation $\hat{g}$ of g in the local occupation number Fock-space

$$\hat{g} = \sum \mathcal{D}_{\Gamma\Gamma'}(g)|\Gamma\rangle\langle\Gamma'|\,. \tag{6.48}$$

If the Gutzwiller wave-function is invariant with respect to the group

$$\hat{g}|\Psi_G\rangle = |\Psi_G\rangle\,, \forall g \in \mathcal{G}\,, \tag{6.49}$$

then by the same argument as for the general operator $\hat{\mathcal{O}}(\mathbf{R})$ the factorization ϕ of $\hat{\rho}$ must commute with the group elements

$$[\phi, \hat{g}] = 0\,, \forall g \in \mathcal{G}\,. \tag{6.50}$$

How to exploit point group symmetries are further discussed in Chapter 7.

6.5 Entanglement entropy

Having direct access to the local many-body density matrix $\hat{\rho}$ through the ϕ operator $\hat{\rho} = \phi\phi^\dagger$, do not only enables the calculation of expectation values of arbitrary local operators by a trace expression as in Eq. (6.41). Another useful quantity is the local entanglement entropy S_E [118] defined as

$$S_E = \mathrm{Tr}[\hat{\rho} \log \hat{\rho}] \,. \tag{6.51}$$

The entanglement entropy is a measure on the entanglement between a subset of a system with the rest. Here the subset is the local many-body space on a lattice site, while the rest of the system is the entire lattice with one lattice point removed.

The extreme limits of S_E is obtained from the non-interacting system at half-filling. In this limit all local many-body states are occupied with equal probability and S_E simplifies to the logarithm of the size of the local Fock-space N_Γ

$$S_E = \log N_\Gamma \,. \tag{6.52}$$

In the atomic limit with no hopping S_E simply counts the degeneracy d of the atomic ground state and gives

$$S_E = \log d \,. \tag{6.53}$$

The entanglement entropy is also sensitive to phase transitions [118] and is a relevant "observable" for mapping out phase diagrams. Papers V and VI employ the entanglement entropy for exactly this purpose.

Apart from being an interesting property to measure, the entanglement entropy has recently been used to extend the Gutzwiller method to finite temperatures [119]. This is a very interesting development, and a multi-band generalization of these ideas certainly deserves attention.

6.6 Vector-space representation

The variational parameter ϕ can practically be represented in terms of the coefficient matrix $\phi_{\Gamma\eta}$ defined in Eq. (6.30), having the dimension of the local Fock-space squared. However, the number of independent coefficients can be vastly reduced by explicitly taking into account the symmetries of the Gutzwiller wave-function. How to practically do this is discussed in Section 6.4, Chapter 7, and Paper III.

For our present purpose it suffices to note that a symmetry analysis yields a set of basis-matrices ϕ_i that are orthonormal in the trace

$$\mathrm{Tr}[\phi_i \phi_j^\dagger] = \delta_{ij}\,, \tag{6.54}$$

and that the most general ϕ obeying these symmetries can be written as a linear combination of the ϕ_i matrices

$$\phi = \sum_i c_i \phi_i\,, \tag{6.55}$$

using a set of complex coefficients c_i. Thus we can replace the matrix-coefficients $\phi_{\Gamma\eta}$ with a coefficient vector $|c\rangle$, and effectively express the local configuration in a well defined vector-space.

To pass from traces over ϕ in Eq. (6.42) to expressions in terms of the coefficient vector $|c\rangle$ we define the matrices

$$\mathcal{U}_{ij} = \mathrm{Tr}[\hat{H}_{loc}\phi_j\phi_i^\dagger]\,, \quad \mathcal{M}_{ij}^{\alpha\beta} = \mathrm{Tr}[\phi_i^\dagger c_\beta^\dagger \phi_j d_\alpha]\,, \quad \mathcal{N}_{ij}^{\alpha\beta} = \mathrm{Tr}[\phi_i^\dagger \phi_j d_\alpha^\dagger d_\beta]\,, \tag{6.56}$$

that act as operators on $|c\rangle$. In this formulation the ϕ trace expressions are replaced with the $|c\rangle$ expectation values

$$\begin{aligned} \mathrm{Tr}[\hat{H}_{loc}\phi\phi^\dagger] &= \langle c|\mathcal{U}|c\rangle\,, \\ \mathrm{Tr}[\phi^\dagger c_\beta^\dagger \phi d_\alpha] &= \langle c|\mathcal{M}^{\alpha\beta}|c\rangle\,, \\ \mathrm{Tr}[\phi^\dagger \phi d_\alpha^\dagger d_\beta] &= \langle c|\mathcal{N}^{\alpha\beta}|c\rangle\,. \end{aligned} \tag{6.57}$$

The minimization problem of Eq. (6.42) expressed in terms of $|c\rangle$ takes the form

$$\begin{aligned} \min_{c,\Psi_0}\{\mathcal{E}(c,\Psi_0) = \langle\Psi_0|\hat{\mathcal{T}}^G(\mathcal{R})|\Psi_0\rangle + \langle c|\mathcal{U}|c\rangle :& \\ \langle\Psi_0|\Psi_0\rangle = 1,\ \langle\Psi_0|d_\alpha^\dagger d_\beta|\Psi_0\rangle = 0, \alpha\neq\beta,& \\ \langle c|c\rangle = 1,\ \langle c|\mathcal{N}^{\alpha\beta}|c\rangle = \langle\Psi_0|d_\alpha^\dagger d_\beta|\Psi_0\rangle\}\,,& \end{aligned} \tag{6.58}$$

where the renormalization matrix $\mathcal{R}$ is given by

$$\langle c|\mathcal{M}^{\alpha\beta}|c\rangle = \sqrt{n_\alpha^0(1-n_\alpha^0)}\mathcal{R}_{\alpha\beta}\,, \quad n_\alpha^0 = \langle\Psi_0|d_\alpha^\dagger d_\alpha|\Psi_0\rangle\,. \tag{6.59}$$

Taking a closer look at the minimization problem in Eq. (6.58) we see that every $|c\rangle$ dependency is quadratic, similar to all persisting explicit Ψ_0 dependencies. The only remaining non-quadratic terms now stems from the renormalization matrix $\mathcal{R}$ in the kinetic term $\hat{\mathcal{T}}^G(\mathcal{R}) \propto \mathcal{R}^2$, where $\mathcal{R}$ in turn depends quadratically on $|c\rangle$, and on Ψ_0, through the square root expression of the local density matrix n^0 due to Eq. (6.59).

6.7 Program separation

Looking at Eqs. (6.58) and (6.59) one realizes that the bothersome direct connection between Ψ_0 and ϕ through $\mathcal{R}$ and n^0 actually can be exploited. The connection can be made implicit by promoting $\mathcal{R}$ and n^0 to free parameters, and using Eq. (6.59) as a constraint instead of a direct relation. This procedure gives the extended minimization problem

$$\begin{aligned}\min_{\Psi_0,c,\mathcal{R},n^0}\{\mathcal{E}(\Psi_0,c,\mathcal{R},n^0) = \langle\Psi_0|\hat{\mathcal{T}}^G(\mathcal{R})|\Psi_0\rangle + \langle c|\mathcal{U}|c\rangle :\\ \langle\Psi_0|\Psi_0\rangle = 1,\ \langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle = \delta_{\alpha\beta}n^0_\alpha,\\ \langle c|c\rangle = 1,\ \langle c|\mathcal{N}^{\alpha\beta}|c\rangle = \delta_{\alpha\beta}n^0_\alpha,\\ \langle c|\mathcal{M}^{\alpha\beta}|c\rangle = \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta}\}\,.\end{aligned} \tag{6.60}$$

At first sight it might seem like the problem only has been complicated, but promoting $\mathcal{R}$ and n^0 to free parameters makes the minimization separable in Ψ_0 and $|c\rangle$. In fact, we can directly write Eq. (6.60) in two steps, as an outer minimization with respect to $\mathcal{R}$ and n^0

$$\min_{\mathcal{R},n^0}\{\mathcal{E}_{\Psi_0}(\mathcal{R},n^0) + \mathcal{E}_c(\mathcal{R},n^0)\}\,. \tag{6.61}$$

of the minimal kinetic energy

$$\begin{aligned}\mathcal{E}_{\Psi_0}(\mathcal{R},n^0) = \min_{\Psi_0}\{\langle\Psi_0|\hat{\mathcal{T}}^G(\mathcal{R})|\Psi_0\rangle : \langle\Psi_0|\Psi_0\rangle = 1,\\ \langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle = \delta_{\alpha\beta}n^0_\alpha\}\,,\end{aligned} \tag{6.62}$$

and the minimal interaction energy

$$\begin{aligned}\mathcal{E}_c(\mathcal{R},n^0) = \min_{c}\{\langle c|\mathcal{U}|c\rangle : \langle c|c\rangle = 1,\\ \langle c|\mathcal{M}^{\alpha\beta}|c\rangle = \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta},\ \langle c|\mathcal{N}^{\alpha\beta}|c\rangle = \delta_{\alpha\beta}n^0_\alpha\}\,,\end{aligned} \tag{6.63}$$

at fixed $\mathcal{R}$ and n^0. The result is two separate inner minimizations in Ψ_0 and $|c\rangle$ [Eqs. (6.62) and (6.63)]. Both these minimization problems are Quadratic Constrained Quadratic Programs (QCQPs), as all target functions and constraints are quadratic forms.

In general QCQPs are non-deterministic polynomial-time hard (i.e. NP-hard), but when the operators are positive semi-definite they become convex problems. Thus, even though finding the global solution is hard in general,

local minima can be computed in a straight forward manner using for example Sequential Least SQuares Programming (SLSQP) [120]. The SLSQP method uses the gradients of both the target and constraint functions, which for a QCQP is a cheap matrix vector multiplication.

We have performed exploratory tests of solving the separated minimization of Eq. (6.61) using the bounded Simplex Algorithm, applying SLSQP for the $|c\rangle$ minimization, with promising results. However this approach is computationally much more costly than the Lagrange based approach that is the subject of next section.

6.8 Lagrange formulation

The alternative to direct minimization of Eqs. (6.62) and (6.63) is to use the Lagrange multiplier method to find stationary feasible points of the energies $\mathcal{E}_{\Psi_0}$ and $\mathcal{E}_c$. The Slater determinant problem Eq. (6.62) has the associated Lagrange function

$$\mathcal{L}_{\Psi_0}(\Psi_0, E_{\Psi_0}, \lambda^{(\Psi_0)}) = \langle\Psi_0|\hat{\mathcal{T}}^G(\mathcal{R})|\Psi_0\rangle + E_{\Psi_0}(1 - \langle\Psi_0|\Psi_0\rangle) + \frac{1}{2}\sum_{\alpha\beta}\left[\lambda^{(\Psi_0)}_{\alpha\beta}(\langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle - \delta_{\alpha\beta}n^0_\alpha) + \text{c.c.}\right], \quad (6.64)$$

where every constraint enters multiplied by one of the Lagrange multipliers E_{Ψ_0} and $\lambda^{(\Psi_0)}$. In writing down the density constraint we have used the relations for complex constraints and Lagrange multipliers discussed in Appendix C. A necessary and sufficient condition for feasible stationary points of $\mathcal{E}_{\Psi_0}$ is that the associated Lagrange function $\mathcal{L}_{\Psi_0}$ is stationary $\nabla\mathcal{L}_{\Psi_0} = \mathbf{0}$ (in all variables Ψ_0, E_{Ψ_0} and $\lambda^{(\Psi_0)}$). In principle it is possible to solve for stationarity directly.

Calculating the explicit partial derivatives with respect to the Lagrange multipliers E_{Ψ_0} and $\lambda^{(\Psi_0)}$ reproduces the constraints of Eq. (6.62)

$$\frac{\partial\mathcal{L}_{\Psi_0}}{\partial E_{\Psi_0}} = 1 - \langle\Psi_0|\Psi_0\rangle = 0\,, \quad (6.65)$$

$$2\frac{\partial\mathcal{L}_{\Psi_0}}{\partial\lambda^{(\Psi_0)}_{\alpha\beta}} = \langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle - \delta_{\alpha\beta}n^0_\alpha = 0\,. \quad (6.66)$$

The partial derivative with respect to $\langle\Psi_0|$ is more intriguing

$$\frac{\partial\mathcal{L}_{\Psi_0}}{\partial\langle\Psi_0|} = \left[\hat{\mathcal{T}}^G(\mathcal{R}) + \frac{1}{2}\sum_{\alpha\beta}\left(\lambda^{(\Psi_0)}_{\alpha\beta}d^\dagger_\alpha d_\beta + \text{h.c.}\right) - E_{\Psi_0}\right]|\Psi_0\rangle = 0\,, \quad (6.67)$$

as it takes the form of an eigenstate problem in $|\Psi_0\rangle$. The ground state gives the minimal constrained kinetic energy, so Ψ_0 is uniquely determined by $\lambda^{(\Psi_0)}$ through Eq. (6.67), $\Psi_0 = \Psi_0(\lambda^{(\Psi_0)})$. This allows us to replace the stationarity problem $\nabla\mathcal{L}_{\Psi_0}(\Psi_0; E_{\Psi_0}, \lambda^{(\Psi_0)}) = \mathbf{0}$ with the root problem $R(\lambda^{(\Psi_0)}) = \mathbf{0}$, where the function $R(\lambda^{(\Psi_0)})$ incorporates the steps:

i) Given $\lambda^{(\Psi_0)}$, compute the ground state eigenvector Ψ_0 of Eq. (6.67).

ii) Normalize Ψ_0 as to fulfill the normalization constraint of Eq. (6.65).

iii) Use Ψ_0 to evaluate the only remaining constraint (Eq. (6.66)) and return the result.

Finding the root of $R(\lambda^{(\Psi_0)}) = \mathbf{0}$ now gives a solution of the higher dimensional problem $\nabla\mathcal{L}_{\Psi_0}(\Psi_0; E_{\Psi_0}, \lambda^{(\Psi_0)}) = \mathbf{0}$. The great benefit of this embedding of Eqs. (6.65) and (6.67) in $R(\lambda^{(\Psi_0)})$ is the reduction in dimensionality of the root problem that in the end has to be solved using a Newton based numerical method.

Applying the same procedure to the minimization problem with respect to $|c\rangle$ in Eq. (6.63) gives the associated Lagrange function

$$\begin{aligned}\mathcal{L}_c(c; E_c, \lambda^{(c)}, \Lambda) = \langle c|\mathcal{U}|c\rangle + E_c(1 - \langle c|c\rangle) \\ + \frac{1}{2}\sum_{\alpha\beta}\left[\lambda^{(c)}_{\alpha\beta}(\langle c|\mathcal{N}^{\alpha\beta}|c\rangle - \delta_{\alpha\beta}n^0_\alpha) + \text{c.c}\right] \\ + \frac{1}{2}\sum_{\alpha\beta}\left[\Lambda_{\alpha\beta}(\langle c|\mathcal{M}^{\alpha\beta}|c\rangle - \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta}) + \text{c.c}\right], \quad (6.68)\end{aligned}$$

where we, apart from the normalization and density constraints, have introduced extra Lagrange multipliers Λ to fulfill the constraint on $|c\rangle$ with respect to the renormalization matrix $\mathcal{R}$. We also here seek to solve for stationary points to the Lagrange function $\nabla\mathcal{L}_c = \mathbf{0}$. The Lagrange multiplier partial derivatives again reproduce the constraints of the minimization problem [Eq. (6.63)]

$$\frac{\partial\mathcal{L}_c}{\partial E_c} = 1 - \langle c|c\rangle = 0\,, \quad (6.69)$$

$$2\frac{\partial\mathcal{L}_c}{\partial\lambda^{(c)}_{\alpha\beta}} = \langle c|\mathcal{N}^{\alpha\beta}|c\rangle - \delta_{\alpha\beta}n^0_\alpha = 0\,, \quad (6.70)$$

$$2\frac{\partial\mathcal{L}_c}{\partial\Lambda_{\alpha\beta}} = \langle c|\mathcal{M}^{\alpha\beta}|c\rangle - \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta} = 0\,. \quad (6.71)$$

The partial derivative with respect to $\langle c|$ also takes the form of the eigenstate problem

$$\frac{\partial \mathcal{L}_c}{\partial \langle c|} = \left[\mathcal{U} + \frac{1}{2} \sum_{\alpha\beta} \left(\lambda^{(c)}_{\alpha\beta} \mathcal{N}^{\alpha\beta} + \Lambda_{\alpha\beta} \mathcal{M}^{\alpha\beta} + \text{h.c.} \right) - E_c \right] |c\rangle = 0 \,, \quad (6.72)$$

where $\lambda^{(c)}$ and Λ enters as "single-particle" potentials. In analogue with the Slater determinant Ψ_0 case, the ground state $|c\rangle$ of Eq. (6.72) is completely determined by the Lagrange multipliers $\lambda^{(c)}$ and Λ, $|c\rangle = |c(\lambda^{(c)}, \Lambda)\rangle$. Thus the stationarity problem in $\mathcal{L}_c$ can be replaced by a reduced root problem, along the same lines as outlined for Ψ_0.

The previous treatments of the Ψ_0 and $|c\rangle$ optimizations where performed under the assumption of fixed n^0 and $\mathcal{R}$. However, this is not necessary and we are now going to lift this limitation and reformulate also the minimization with respect to n^0 and $\mathcal{R}$ for fixed total density N. To this end we look for stationary solutions to the associated Lagrange function $\mathcal{L}$ of the full problem in Eq. (6.60)

$$\begin{aligned} &\mathcal{L}(\Psi_0, c; \mathcal{R}, n^0; E_{\Psi_0}, E_c; \lambda^{(\Psi_0)}, \lambda^{(c)}, \Lambda; \mu) \\ &\quad = \mathcal{L}_{\Psi_0}(\Psi_0; E_{\Psi_0}, \lambda^{(\Psi_0)}) + \mathcal{L}_c(c; E_c, \lambda^{(c)}, \Lambda) + \mu \left(\sum_\alpha n^0_\alpha - N \right) , \end{aligned} \quad (6.73)$$

where $\mathcal{L}$ is simply the combination of $\mathcal{L}_{\Psi_0}$, $\mathcal{L}_c$ and the Lagrange multiplier term for the total density constraint. The promotion of n^0 and $\mathcal{R}$ to free variables and the introduction of the chemical potential μ now gives the extra partial derivatives

$$\frac{\partial \mathcal{L}}{\partial \mu} = \sum_\alpha n^0_\alpha - N = 0 \,, \quad (6.74)$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{R}_{\alpha\beta}} = \langle \Psi_0 | \frac{\partial \hat{\mathcal{T}}^G(\mathcal{R})}{\partial \mathcal{R}_{\alpha\beta}} | \Psi_0 \rangle + \frac{1}{2} \sqrt{n^0_\alpha (1 - n^0_\alpha)} \Lambda_{\alpha\beta} = 0 \,, \quad (6.75)$$

$$-\frac{\partial \mathcal{L}}{\partial n^0_\alpha} = \frac{1}{2} \frac{1 - 2n^0_\alpha}{\sqrt{n^0_\alpha (1 - n^0_\alpha)}} \frac{1}{2} \sum_\beta (\Lambda_{\alpha\beta} \mathcal{R}_{\alpha\beta} + \text{c.c.}) + \lambda^{(\Psi_0)}_{\alpha\alpha} + \lambda^{(c)}_{\alpha\alpha} - \mu = 0 \,, \quad (6.76)$$

while the other partial derivatives of $\mathcal{L}$ still are given by Eqs. (6.65), (6.66), (6.67), (6.69), (6.70), (6.71) and (6.72).

We can of course still benefit from the mappings $\lambda^{(\Psi_0)} \rightarrow \Psi_0$ and $(\lambda^{(c)}, \Lambda) \rightarrow |c\rangle$ by constructing a root function R in a reduced set of variables. Just as was done in the separated minimization for $\mathcal{E}_{\Psi_0}$ and $\mathcal{E}_c$. Let $R(\mathcal{R}, \lambda^{(\Psi_0)}, \lambda^{(c)}_{\alpha\neq\beta}, \mu)$ and take the steps

i) Given $\mathcal{R}$ and $\lambda^{(\Psi_0)}$, compute the ground state eigenvector Ψ_0 of Eq. (6.67).

ii) Use Ψ_0 and the diagonal part of Eq. (6.66) to calculate the density matrix in the natural basis $n^0_\alpha = \langle\Psi_0|d^\dagger_\alpha d_\alpha|\Psi_0\rangle$.

iii) Use $\mathcal{R}$, Ψ_0, and n^0 to compute Λ using Eq. (6.75).

iv) Use $\mathcal{R}$, n^0, $\lambda^{(\Psi_0)}$, μ, and Λ to compute $\lambda^{(c)}_{\alpha\alpha}$ through Eq. (6.76).

v) Use the now complete $\lambda^{(c)}$ and Λ to compute $|c\rangle$ as the ground state of Eq. (6.72).

vi) With all parameters known, evaluate the left hand side of all Lagrange partial derivative equations [Eqs. (6.66), (6.70), (6.71) and (6.74)] that have not previously been used

$$\begin{aligned}
\langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle &= 0\,, \text{ for } \alpha \neq \beta\,,\\
\langle c|\mathcal{N}^{\alpha\beta}|c\rangle - \delta_{\alpha\beta} n^0_\alpha &= 0\,,\\
\langle c|\mathcal{M}^{\alpha\beta}|c\rangle - \sqrt{n^0_\alpha(1-n^0_\alpha)}\mathcal{R}_{\alpha\beta} &= 0\,,\\
\sum_\alpha n^0_\alpha - N &= 0\,,
\end{aligned}$$

and return the result.

Thus solving for $R(\mathcal{R}, \lambda^{(\Psi_0)}, \lambda^{(c)}_{\alpha\neq\beta}, \mu) = \mathbf{0}$ ensures that the solution is a stationary point of the full Lagrange function $\mathcal{L}$.

In most applications in this thesis the Hamiltonians under study have diagonal single-particle density matrices by construction. This makes the original basis $c^\dagger_\alpha$ equal to the Slater determinant natural basis $d^\dagger_\alpha = c^\dagger_\alpha$ and guarantees that

$$\langle\Psi_0|c^\dagger_\alpha c_\beta|\Psi_0\rangle = \langle\Psi_0|d^\dagger_\alpha d_\beta|\Psi_0\rangle\,. \tag{6.77}$$

In this case there is no need for off-diagonal Lagrange multipliers to impose diagonality of the single-particle density matrices, and it is enough to work only with $\lambda^{(\Psi_0)}_{\alpha\alpha}$ and $\lambda^{(c)}_{\alpha\alpha}$.

Chapter 7

Point Group Symmetry

This chapter describes how lattice point group symmetries can be incorporated in local many-body operators. The basis of symmetric operators is considerably reduced compared to the most general basis, and the reduction has been essential for making Gutzwiller calculations for d-electron models computationally feasible. As the theory of finite groups is text-book material, theorems and lemmas are simply stated, omitting proofs. The interested reader is encouraged to indulge in Refs. [41] and [121], which have inspired this summary.

The discussion is organized as follows; Section 7.1 briefly introduces finite groups, irreducible representations, and the Schur lemma. The concept of character vectors and Dirac characters are developed in Section 7.2 and used to analyze the irreducible content of a general representation. The great orthogonality theorem[1] is also presented and used to find the blocking transform of a representation. The general discussion on group theory is closed by Section 7.3, where we develop a general scheme for constructing all irreducible representations of a group from any faithful representation. Section 7.4 introduces proper lattice point groups and the two explicit examples of the cubic and tetragonal point groups. The angular momentum representation is presented in Section 7.5, and used in Section 7.6 to construct a basis for point group invariant many-body operators.

[1] One of the most versatile tools in the group theory toolbox.

7.1 Finite groups

Consider a general finite group $\mathcal{G}$ with group elements g. Every homomorphic mapping of g to non-singular matrices $R(g)$ with dimensions $d \times d$ is then a d dimensional *representation* $R(\mathcal{G})$ of $\mathcal{G}$. For finite groups the representation can be chosen to be unitary without any loss of generality, as this merely corresponds to the choice of an orthogonal basis [41]. Two representations $R(g)$ and $R'(g)$ are said to be equivalent $R \approx R'$, if they are related by an unitary transform $\mathcal{U}$

$$R'(g) = \mathcal{U}^\dagger R(g) \mathcal{U}\,, \quad \forall g \in \mathcal{G}\,. \tag{7.1}$$

A representation R is said to be *reducible* if it can be block-diagonalized

$$\mathcal{U}^\dagger R(g) \mathcal{U} = \begin{bmatrix} R^{(1)}(g) & 0 & \dots & 0 \\ 0 & R^{(2)}(g) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & R^{(n)}(g) \end{bmatrix}, \forall g \in \mathcal{G}\,, \tag{7.2}$$

by an unitary transform $\mathcal{U}$, which we will write in shorthand as

$$\mathcal{U}^\dagger R(g) \mathcal{U} = \bigoplus_{i=1}^{n} R^{(i)}(g)\,. \tag{7.3}$$

Conversely, if a representation R is not reducible by any unitary transform then it is *irreducible*. Thus, a reducible representation R can always be block-diagonalized in blocks of irreducible representations $R^{(i)}$ as indicated in Eq. (7.2).

The special properties of the block-diagonal form is described by the Schur lemma [41]. Let R be a reducible representation that can be brought to block-diagonal form with n irreducible representations $R^{(i)}$ each one repeated r_i times, i.e. R can be written as

$$\mathcal{U}^\dagger R(g) \mathcal{U} = \bigoplus_{i=1}^{n} \left(\bigoplus_{r=1}^{r_i} R^{(i)}(g) \right)\,. \tag{7.4}$$

The Schur lemma now states that the most general matrix B that commutes with all group elements $[B, R(g)] = 0$, $\forall g \in \mathcal{G}$, has the form

$$\mathcal{U}^\dagger B \mathcal{U} = \bigoplus_{i=1}^{n} \left(b^{(i)} \otimes \mathbf{1}_{d_i} \right)\,, \tag{7.5}$$

where d_i is the dimension of $R^{(i)}$, $\mathbf{1}_{d_i}$ is the d_i-dimensional identity matrix, $b^{(i)}$ are matrices with arbitrary coefficients and dimension $r_i \times r_i$, and the Kronecker product gives

$$b^{(i)} \otimes \mathbf{1}_{d_i} = \begin{bmatrix} b^{(i)}_{11}\mathbf{1}_{d_i} & \dots & b^{(i)}_{1r_i}\mathbf{1}_{d_i} \\ \vdots & \ddots & \vdots \\ b^{(i)}_{r_i1}\mathbf{1}_{d_i} & \dots & b^{(i)}_{r_ir_i}\mathbf{1}_{d_i} \end{bmatrix} . \tag{7.6}$$

Thus, the number of independent parameters $n_b = \sum_{i=1}^{n} r_i^2$ in $b^{(i)}$, producing matrices B invariant under $R(\mathcal{G})$ is considerably lower than the total number of elements $N_B = \left(\sum_{i=1}^{n} d_i r_i\right)^2$ in B.

The space of matrices B is isomorphic to the vector space of coefficients c_η with respect to a set of basis matrices B_η

$$B = \sum_\eta c_\eta B_\eta \,, \tag{7.7}$$

where B_η can be directly constructed from the general invariant expression in Eq. (7.5) by letting $b^{(i)}$ have only one non-zero coefficient

$$\mathcal{U}^\dagger B_\eta \mathcal{U} \equiv \mathcal{U}^\dagger B^{(j)}(rr')\mathcal{U} = \bigoplus_{i=1}^{n_i} \left(b^{(ij)}_{rr'} \otimes \mathbf{1}_{d_i} \right) , \tag{7.8}$$

where η is an enumeration of all possible elements in $b^{(i)}$, $\eta = (j, rr')$, and $b^{(ij)}_{rr'} = \frac{1}{\sqrt{d_i}}\delta_{ij}\hat{e}^{(r_i)}_r \cdot (\hat{e}^{(r_i)}_{r'})^T$. Here $\hat{e}^{(r_i)}_r$ is the r_i dimensional unit vector pointing in dimension r, and $\hat{e}^{(r_i)}_r \cdot (\hat{e}^{(r_i)}_{r'})^T$ is a $r_i \times r_i$ matrix with the element in row r and column r' equal to one and zero everywhere else.

By this choice the basis matrices B_η are orthonormal with respect to the trace

$$\mathrm{Tr}[B^\dagger_\eta B_\xi] = \delta_{\eta\xi} \,, \tag{7.9}$$

and the coefficients c_η of a general matrix are given by

$$c_\eta = \mathrm{Tr}[B^\dagger_\eta B] \,. \tag{7.10}$$

Thus, imposing invariance of a matrix with respect to a representation of a finite group $\mathcal{G}$, amounts to finding the unitary transform $\mathcal{U}$ and using the Schur lemma. However, determining $\mathcal{U}$ is non-trivial.

7.2 Character vectors and Dirac characters

There are more than one way to find the blocking transform of a representation. Here we will present a method directly based on group theory considerations.

The irreducible representations $R^{(i)}$ are intimately connected with the conjugacy classes $\mathcal{C}$ of a group $\mathcal{G}$, which are defined as the subsets of elements $g \in \mathcal{G}$ related by the conjugacy transform

$$g_1, g_2 \in \mathcal{C} \iff \exists g \in \mathcal{G} : g_1 = g^{-1} g_2 g \,. \tag{7.11}$$

The number of classes n_c in $\mathcal{G}$ is equal to the number of irreducible representations $n = n_c$, and there are several class properties that can be used to determine the decomposition into irreducible representations.

One such class property is the character χ of a representation R, defined as the trace of $R(g)$

$$\chi(g) = \mathrm{Tr}[R(g)] = \chi(\mathcal{C}) \,, g \in \mathcal{C} \,, \tag{7.12}$$

where $\chi(g)$ is equal for elements in the same class, due to the invariance of the trace with respect to a unitary transform.

The character vector χ for all classes in $\mathcal{G}$ uniquely determines the irreducible representations contained in a representation R. The occurrence and number of repetitions r_i of an irreducible representation $R^{(i)}$ in R is given by the vector dot product between the character vectors of the two representations $\chi^{(i)}$ and χ respectively

$$r_i = \chi \cdot \chi^{(i)} = \frac{1}{n_g} \sum_{g \in \mathcal{G}} \chi(g) \chi^{(i)}(g)^* = \frac{1}{n_c} \sum_{\mathcal{C}} \chi(\mathcal{C}) \chi^{(i)}(\mathcal{C})^* \,. \tag{7.13}$$

As a consequence, a sufficient and necessary condition for a representation to be irreducible is that the norm of its character vector is unity

$$\chi^{(i)} \cdot \chi^{(i)} = \frac{1}{n_g} \sum_{g \in \mathcal{G}} |\chi^{(i)}(g)|^2 = \frac{1}{n_c} \sum_{\mathcal{C}} |\chi^{(i)}(\mathcal{C})|^2 = 1 \,. \tag{7.14}$$

Thus, the decomposition of a representation R as given in Eq. (7.4) can unambiguously be determined from its character vector χ, but it does not give the form of the blocking unitary transform $\mathcal{U}$.

Another property of the classes that gives a partial description of $\mathcal{U}$ are the matrices $\Omega_{\mathcal{C}}$, constructed as sums over all elements in each class $\mathcal{C}$

$$\Omega_{\mathcal{C}} = \frac{1}{n_g} \sum_{g \in \mathcal{G}} R^\dagger(g) R(g') R(g) = \frac{1}{n_c} \sum_{g \in \mathcal{C}} R(g) \,, \quad g' \in \mathcal{C} \,. \tag{7.15}$$

The matrices $\Omega_{\mathcal{C}}$ are called the Dirac characters of R. By the class rearrangement theorem [121], they commute with all group elements $[\Omega_{\mathcal{C}}, R(g)] = 0$, and among themselves $[\Omega_{\mathcal{C}}, \Omega_{\mathcal{C}'}] = 0$. In other words, the Dirac characters are the finite group equivalent to the Casimir operators of continuous Lie groups. Simultaneous diagonalization of all $\Omega_{\mathcal{C}}$ gives a transform $\mathcal{U}_D$ that diagonalizes R in blocks for each irreducible representation $\tilde{R}^{(i)}$.

However, $\mathcal{U}_D$ can not separate repeated irreducible representations $\tilde{R}^{(i)}$, as the eigenvalues of $\Omega_{\mathcal{C}}$ are the same for every pair of equivalent irreducible representations $\tilde{R}^{(i)} = \mathcal{U}^\dagger R^{(i)} \mathcal{U}$. Therefore, the particular form of the irreducible representations $\tilde{R}^{(i)}$ produced by $\mathcal{U}_D$ is determined by the applied diagonalization algorithm, due to the invariance with respect to similarity transforms of $\tilde{R}^{(i)}$.

Fortunately, the great orthogonality theorem can be used to determine a transformation that separates all repeated irreducible representations and produces the blocked form in terms of a specific set of known irreducible representations $R^{(i)}$, as described by Koster in Ref. [122]. The part $\mathcal{U}_r^{(i)}$ of the blocking transform $\mathcal{U}$ corresponding to the matrix-block of the rth repetition of the irreducible representation $R^{(i)}$ in R can be obtained from

$$\frac{1}{n_g} \sum_{g \in G} R(g)_{\alpha\beta} R^{(i)}(g)_{\gamma\delta} = (\mathcal{U}_r^{(i)})^*_{\alpha\delta} (\mathcal{U}_r^{(i)})_{\beta\delta} \,, \tag{7.16}$$

by finding a fixed pair (α, δ) for which $(\mathcal{U}_r^{(i)})_{\alpha\delta} \neq 0$. This procedure gives $\mathcal{U}_r^{(i)}$ up to a constant factor, fixed by a subsequent normalization. When $R^{(i)}$ is repeated r_i times in R there are r_i different pairs (α, δ) that give the transform $\mathcal{U}_r^{(i)}$ for every repetition.

The obtained matrices can be collected into the complete transform matrix

$$\mathcal{U} = \left[\begin{array}{cccccccc} \mathcal{U}_1^{(1)} & \dots & \mathcal{U}_{r_1}^{(1)} & \mathcal{U}_1^{(2)} & \dots & \mathcal{U}_{r_{n_i}}^{(n)} \end{array} \right] , \tag{7.17}$$

that cast R in to the block form of Eq. (7.4)

$$\mathcal{U}^\dagger R(g) \mathcal{U} = \bigoplus_{i=1}^{n} \left(\bigoplus_{r=1}^{r_i} R^{(i)}(g) \right) . \tag{7.18}$$

The above procedure is a complete prescription for finding the blocking transform $\mathcal{U}$ of a representation R when a complete set of irreducible representations $R^{(i)}$ is known beforehand. Thus, we are left with the problem to construct all irreducible representations $R^{(i)}$ of the group $\mathcal{G}$.

7.3 The regular representation

All properties of a finite group are determined by the result of the group element multiplications

$$g_i^{-1} g_j = g_k \,. \tag{7.19}$$

described by the group multiplication table. By enumerating all group elements $g_1, \ldots, g_{n_g}$, the multiplication table can be written as a matrix with integer entries $M_{ij} = k$ specified by the result of the multiplication in Eq. (7.19).

The multiplication table M can be constructed from any faithful representation R_f that is isomorphic with $\mathcal{G}$, and can in turn be used to construct a special representation $\mathcal{R}$ called the *regular representation*, containing all irreducible representations if $\mathcal{G}$. The elements of $\mathcal{R}(g_l)$ are directly obtained from M by setting all entries to zero except those corresponding to the particular element g_l, which instead are set to one

$$\mathcal{R}(g_l) = \delta(M_{ij} - l) \,. \tag{7.20}$$

The regular representation $\mathcal{R}$ is reducible and contains all irreducible representations $R^{(i)}$ repeated the same number of times r_i as the dimension d_i of the irreducible representation $r_i = d_i$, i.e.

$$\mathcal{U}^\dagger \mathcal{R}(g) \mathcal{U} = \bigoplus_{i=1}^{n_i} \left(\bigoplus_{r=1}^{d_i} R^{(i)}(g) \right) \,. \tag{7.21}$$

The Dirac characters of $\mathcal{R}$ are real and symmetric, and have a real blocking transform $\mathcal{U}_D$, that gives the blocks of repeated irreducible representations the particular internal block structure

$$\begin{bmatrix} A_0 & A_1 \\ A_1 & A_0 \end{bmatrix}, \quad \begin{bmatrix} A_0 & A_1 & A_2 \\ A_2 & A_0 & A_1 \\ A_1 & A_2 & A_0 \end{bmatrix}, \quad \begin{bmatrix} A_0 & A_1 & A_2 & A_3 \\ A_3 & A_0 & A_1 & A_2 \\ A_2 & A_3 & A_0 & A_1 \\ A_1 & A_2 & A_3 & A_0 \end{bmatrix}, \quad \text{etc.}, \tag{7.22}$$

where the block-matrices have dimensions $r_i \times r_i$, the blocks A_i dimensions $d_i \times d_i$, and every row in the block-matrix is a cyclic shift of the previous row. It turns out that the block structure can always be put on diagonal form, and that there is only one corresponding real eigenvector. This eigenvector is the "unity" vector with a factor one in every element, and it can be used

to construct the part $\mathcal{U}^{(i)}$ of the blocking-transform

$$\mathcal{U}^{(i)} = \frac{1}{\sqrt{d_i}} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \otimes \mathbf{1}_{d_i} \,, \tag{7.23}$$

that gives a real representation of the irreducible representation $R^{(i)}$. So collecting $\mathcal{U}^{(i)}$ in

$$\tilde{\mathcal{U}} = \bigoplus_{i=1}^{n_i} \mathcal{U}^{(i)} \,, \tag{7.24}$$

gives the real irreducible representations $R^{(i)}$ of $\mathcal{G}$ through

$$\tilde{\mathcal{U}}^{\dagger} \mathcal{U}_D^{\dagger} \mathcal{R}(g) \mathcal{U}_D \tilde{\mathcal{U}} = \bigoplus_{i=1}^{n_i} R^{(i)}(g) \,. \tag{7.25}$$

Combining this scheme with the previous section, any representation R can be decomposed in irreducible representations $R^{(i)}$, as long as we have one faithful representation R_f of the group as starting point. The necessary steps are then:

i) Compute the multiplication table M from R_f.

ii) Construct the regular representation $\mathcal{R}$.

iii) Determine all irreducible representations $R^{(i)}$ from $\mathcal{R}$.

iv) Use the character theorems to determine the irreducible content of R.

v) Block R in irreducible representations using the Dirac characters $\Omega_{\mathcal{C}}$.

vi) Use the great orthogonality theorem to fix the representations and to separate repeated irreducible representations.

7.4 Lattice point groups

Let us now narrow the discussion down to the case of $\mathcal{G}$ being a proper lattice point group, defined as the finite group of proper rotations leaving a periodic lattice unchanged. The defining representation for $\mathcal{G}$ is a sub-group of the special orthogonal Lie group $SO(3)$, i.e. the group of proper 3×3 real space rotation matrices. The reason for not considering standard lattice point

groups, which also includes inversion and improper rotations, is that the spherical harmonics has an inversion symmetry (with parity $(-1)^L$) [121].

For a given atomic structure there are today algorithms for automatic symmetry analysis that determine the point group, and returns its representation in $SO(3)$. One freely available implementation is the SPGlib library [123]. For a given point group the representation can also be obtained from the Atomic Simulation Environment (ASE) [124]. As the $SO(3)$ representation of $\mathcal{G}$ is faithful, the procedure outlined in the previous section can be used to determine all irreducible representations. However, it is of course also possible to refer to standard tables over point groups, e.g. the Bilbao Crystallographic Server [125].

As an example, the irreducible representations and their character vectors for all classes of the cubic point group O (432) and the tetragonal point group D_4 (422) are shown in Tables 7.1 and 7.2, respectively. Note that the number of classes and irreducible representations are the same. The dimension of each irreducible representation can be directly read of from the character of the unit element E, which forms a class on its own. The other classes contain n-fold rotations denoted by NC_n, where N is the number of class elements.

The knowledge of the $SO(3)$ rotation matrix representation and all irreducible representations $R^{(i)}$ can now be combined with the decomposition scheme in Section 7.2 and the Schur lemma from Section 7.1 to construct matrices invariant with respect to the point group. More specifically, to construct all point group invariant operators in the local many-body basis from the total angular momentum basis. But to accomplish this, we first need the representation of the point group rotations in angular momentum space.

However, before treating angular momentum we first consider two alternative ways of representing real space rotations. The defining representation $SO(3)$ gives the group elements g in terms of 3×3 orthogonal matrices $R(g)$. An alternative to this is to use a rotation axis $\vec{n}$ and a rotation angle θ combined in the versor $\vec{\theta} = \theta\vec{n}$, which gives the $SO(3)$ rotation matrix representation as [41]

$$R(g) = e^{-i\vec{\theta}(g)\cdot\mathbf{l}}, \tag{7.26}$$

where the components of $\mathbf{l} = [l_x, l_y, l_z] = [i\epsilon_{1ij}, i\epsilon_{2ij}, i\epsilon_{3ij}]$, are the real space generators of rotations, and ϵ_{kij} is the Levi-Civita tensor. Another useful form for rotations is the Euler angles α, β, γ, which define a series of three rotations around specific axes

$$R(g) = e^{-i\alpha l_z} e^{-i\beta l_y} e^{-i\gamma l_z}. \tag{7.27}$$

Table 7.1: Character table for the cubic point group O (432).

Irrep \ Class	E	$6C_4$	$8C_3$	$6C_2$	$3C_2$
A_1	1	1	1	1	1
A_2	1	-1	1	-1	1
E_1	2	0	-1	0	2
T_1	3	1	0	-1	-1
T_2	3	-1	0	1	-1

Table 7.2: Character table for the tetragonal point group D_4 (422).

Irrep \ Class	E	$2C_4$	$2C_2$	$2C_2$	$1C_2$
A_1	1	1	1	1	1
A_2	1	1	-1	-1	1
A_3	1	-1	1	-1	1
A_4	1	-1	-1	1	1
E_1	2	0	0	0	-2

The representation of the rotations $R(g)$ in some other basis can now be obtained by simply replacing $\mathbf{l}$ with the generators of rotation in the particular space of interest.

7.5 Angular momentum representation

The generators of rotations L_i in the angular momentum basis with total angular momentum L have the same form as their second quantized expressions in Eq. (3.8)

$$(L_z)_{MM'} = M\delta_{MM'}\,, \tag{7.28}$$

$$(L_\pm)_{MM'} = \sqrt{L(L+1) - M(M\pm 1)}\delta_{M\pm 1,M'}\,. \tag{7.29}$$

Using the versor $\vec{\theta}(g)$ of a rotation $R(g)$ the corresponding angular momentum representation $\mathcal{D}^{(L)}$ can be obtained by

$$\mathcal{D}^{(L)}(g) = e^{-i\vec{\theta}(g)\cdot\mathbf{L}}\,, \tag{7.30}$$

with $\mathbf{L} = [L_x, L_y, L_z]$ and $L_x = (L_+ + L_-)/\sqrt{2}$ and $L_y = (L_+ - L_-)/(\sqrt{2}i)$. For every L this gives a $(2L+1)$-dimensional representation $\mathcal{D}^{(L)}$ of the point group $\mathcal{G}$. Instead of resorting to compute matrix exponents to determine $\mathcal{D}^{(L)}(g)$, the Wigner functions give a direct algebraic expression for the corresponding rotation in terms of the Euler angles

$$\begin{aligned}\mathcal{D}^{(L)}_{MM'}(\alpha,\beta,\gamma) &= e^{-i\alpha L_z}e^{-i\beta L_y}e^{-i\gamma L_z} \\ &= e^{-i(M\alpha+M'\gamma)}e^{-i\beta L_y} = e^{-i(M\alpha+M'\gamma)}d^{(L)}_{MM'}(\beta)\,,\end{aligned} \tag{7.31}$$

where the function $d^{(L)}_{MM'}$ is called Wigner's formula and has the form [112]

$$\begin{aligned}d^{(L)}_{MM'}(\beta) = \sum_k \frac{\sqrt{(L+M')!(L-M')!(L+M)!(L-M)!}}{(L+M'-k)!k!(L-k-M)!(k-M'+M)!} \\ \times(-1)^{k-M'+M}\left(\cos\frac{\beta}{2}\right)^{2L-2k+M'-M}\left(\sin\frac{\beta}{2}\right)^{2k-M'+M}\,,\end{aligned} \tag{7.32}$$

where the sum is taken over all k for which the denominator is positive. The $\mathcal{D}^{(L)}$ representation is readily decomposed in irreducible representations using the scheme of Section 7.2 giving the blocking transform

$$\mathcal{U}_L^\dagger\mathcal{D}^{(L)}(g)\mathcal{U}_L = \bigoplus_{i\in L}\left(\bigoplus_{r=1}^{r_i^{(L)}} R^{(i)}(g)\right)\,, \tag{7.33}$$

where i is summed over the irreducible representations $R^{(i)}$ contained in $\mathcal{D}^{(L)}$, repeated $r_i^{(L)}$ times. For the cubic point group O (432) and the tetragonal point group D_4 (422) this produces the decompositions listed in Tables 7.3 and 7.4, for the relevant total angular momentum states L of all d-electron many-body states.

Now for every fixed number N of particles we get the complete representation of the point group as

$$\mathcal{D}^{(N)}(g) = \bigoplus_{L \in N} \mathcal{D}^{(L)}(g)\,, \tag{7.34}$$

which is readily decomposed in the irreducible representations $R^{(i)}$ by the blocking transform $\mathcal{U}_N$

$$\mathcal{U}_N^\dagger \mathcal{D}^{(N)}(g) \mathcal{U}_N = \bigoplus_{i \in N} \left(\bigoplus_{r=1}^{r_i} R^{(i)}(g) \right) . \tag{7.35}$$

Finally, applying the Schur lemma from Section 7.1 gives a basis for all point group invariant matrices in the space of N particles as

$$B_\eta^{(N)} = \mathcal{U}_N \left[\bigoplus_{i \in N} \left(b_{rr'}^{(ij)} \otimes \mathbf{1}_{d_i} \right) \right] \mathcal{U}_N^\dagger\,. \tag{7.36}$$

The basis matrices $B_\eta^{(N)}$ can in turn be used to construct all local point group invariant many-body operators.

7.6 Invariant many-body operators

The point group invariant local many-body operators for d-electrons can be expressed directly using the total number N, spin S, angular momentum L, and seniority ν basis $|NS\Sigma LM\nu\rangle$, defined in Section 3.5. A basis for the most general spin rotational and angular momentum point group invariant operator is given by

$$\phi_k = \frac{1}{\sqrt{2S+1}} \sum_{\Sigma} \sum_{LM,L'M'} |NS\Sigma LM\nu\rangle \left(B_\eta^{(N)} \right)_{LM,L'M'} \langle NS\Sigma L'M'\nu'| \tag{7.37}$$

where $k = (N, S, \eta, \nu, \nu')$. The spin rotational invariance is ensured by summing over the diagonal spin components and the point group invariance

Table 7.3: Decomposition of the angular momentum L representations $\mathcal{D}^{(L)}$ in irreducible representations for the cubic point group O (432).

L	$\mathcal{D}^{(L)} \approx \bigoplus R^{(i)}$								
S	A_1								
P	T_1								
D	E_1	$\oplus$	T_2						
F	A_2	$\oplus$	T_1	$\oplus$	T_2				
G	A_1	$\oplus$	E_1	$\oplus$	T_1	$\oplus$	T_2		
H	E_1	$\oplus$	$2T_1$	$\oplus$	T_2				
I	A_1	$\oplus$	A_2	$\oplus$	E_1	$\oplus$	T_1	$\oplus$	$2T_2$

Table 7.4: Decomposition of the angular momentum L representations $\mathcal{D}^{(L)}$ in irreducible representations for the tetragonal point group D_4 (422).

L	$\mathcal{D}^{(L)} \approx \bigoplus R^{(i)}$								
S	A_1								
P	A_2	$\oplus$	E_1						
D	A_1	$\oplus$	A_3	$\oplus$	A_4	$\oplus$	E_1		
F	A_2	$\oplus$	A_3	$\oplus$	A_4	$\oplus$	$2E_1$		
G	$2A_1$	$\oplus$	A_2	$\oplus$	A_3	$\oplus$	A_4	$\oplus$	$2E_1$
H	A_1	$\oplus$	$2A_2$	$\oplus$	A_3	$\oplus$	A_4	$\oplus$	$3E_1$
I	$2A_1$	$\oplus$	A_2	$\oplus$	$2A_3$	$\oplus$	$2A_4$	$\oplus$	$3E_1$

is imposed by using the invariant matrices $B_\eta^{(N)}$, constructed in the previous section.

The many-body operators ϕ_k now span the space of all angular momentum point group and spin rotationally invariant operators, and can directly be used to parametrize the ϕ-operator of the Gutzwiller formalism in terms of complex coefficients c_k

$$\phi = \sum_k c_k \phi_k \,, \tag{7.38}$$

as outlined in Section 6.6.

The use of a point group symmetric basis for many-body operators is essential to reduce the number of free parameters in a Gutzwiller calculation. As an example the d-electron model has roughly a million of free parameters (the Fock-space dimension squared) if one does not consider any symmetry, to be compared to the 873 parameters of a cubic point group symmetric paramagnetic system. The point group symmetric basis construction was used in Paper III to enable bench-mark calculations of d-electrons in cubic crystal fields. The analogous tetragonal construction was later used in Paper IV for a realistic electron structure calculation of the parent states of the iron chalcogenide anomalous superconductors.

Chapter 8

Summary of papers

This chapter gives brief summaries of the appended papers and how they relate to the introductory text of this thesis. The goal is to complement rather than replace the abstracts of the papers in question.

Paper I – Dynamical mean field theory phase-space extension and critical properties of the finite temperature Mott transition

The subject of study in Paper I is the finite temperature Mott metal-insulator transition, in the paramagnetic state of the single-band Hubbard model, on the infinite dimensional Bethe lattice. The transition bears many similarities with the Mott transition in Cr-doped V_2O_3 mentioned in Section 1.2, with a first order line ending in a critical end point. This feature is contained in the paramagnetic state of the Hubbard model as outlined in Section 2.5, where also the Bethe lattice and the limit of infinite dimensions where treated.

In the paper we apply Dynamical Mean Field Theory, introduced in Chapter 4, and the impurity solvers Exact Diagonalization and Iterated Perturbation Theory, from Chapter 5, to study the critical properties of the end point. For this purpose we perform an analysis of the DMFT fixpoint function and its Jacobian, and develop the phase-space extension method from Section 4.3.3 to map out the DMFT fixpoint solutions. The emerging picture is that, within DMFT, the critical end point is a cusp catastrophe in the parameter space in direct analogy to the fourth order Landau model for first order transitions studied in Section 2.6.

The critical exponents are found to have mean field character with a 1/3

exponent on the dependence of the double occupancy D on the interaction U and temperature T, in agreement with previous theoretical studies. The Maxwell construction is applied to determine the location of the first order transition, which enables us to explicitly show that the critical exponent for D along the first order transition line is $1/2$, in agreement with the experimental result for V_2O_3.

Paper II – Distributional exact diagonalization formalism for quantum impurity models

In Paper II we develop a novel solver for the interacting impurity coupled to a continuous bath, from Chapter 5. The basic idea is to extend the Exact Diagonalization formalism, where the continuous bath is approximated with *a single* Anderson model with a finite number of bath states, by using a representable *ensemble* of finite models to mimic the continuous problem.

The resulting approximation was discussed in Section 5.4, and amounts to replacing cross-correlations between finite models with internal correlations. This is a non-trivial approximation and the results gives surprisingly good agreement with high precision methods like NRG, on the real axis, and CT-QMC on the Matsubara axis. The main advantage of our method is that it can be directly extended to compute real axis spectral functions for multi-band problems, currently out of reach for other real axis methods. Also the cross-correlation approximation is interesting per se and its range of applicability deserves a systematic study.

Paper III – Efficient implementation of the Gutzwiller variational method

In Paper III we develop the vector space representation for the Gutzwiller variational problem presented in Section 6.6, and use it to construct a solver for arbitrary local interactions. The solver is a first incarnation of the Lagrange method presented in Section 6.8, limited to fixed variational density n^0, requiring an extra outer minimization of the total energy.

Also the general idea for how to incorporate point group symmetries is presented and a primitive form of the symmetry algorithm treated in Chapter 7 is used to construct the cubic symmetry basis used in the calculations.

The solver is bench marked on a number of standard models by comparison with slave-boson and DMFT results, with nearly exact agreement in the first case and in good agreement in the metallic state for the former.

Paper IV – Orbital selectivity in Hund's metals: The iron chalcogenides

Paper IV is a study of correlation effects in the iron chalcogenide anomalous super conductors FeSe and FeTe, mentioned in Section 1.2. A realistic low energy model is constructed from a DFT calculation along the lines of Sections 1.1 and 3.1. The local d-band interaction is modeled using the Kanamori interaction introduced in Section 3.6, and the low energy model is then solved using the Gutzwiller solver with a tetragonal symmetric basis constructed using the method of Chapter 7.

The results show that the Hund's rule exchange coupling J plays a central role for driving orbital differentiation in these compounds. At intermediate Hubbard interaction U a cross over in to a strongly correlated metal occur, and the rather strongly correlated chalcogenides are classified as Hund's correlated metals.

Paper V – Local correlation in the d-band, Slater-Condon vs. Kanamori

In Paper V we use the parametrization of the full Slater-Condon form of the d-band Coulomb interaction, in terms of the Kanamori parameters U and J and the multipole factor $1/\gamma$, discussed in Section 3.8, to compare the full interaction with the simplified Kanamori form. Special attention is put on the fillings where the Kanamori model has two accidentally degenerate ground state multiplets. The degeneracy increases the overall ground state degeneracy, which in turn modifies the paramagnetic phase diagram.

In the paper we also present the Gutzwiller Lagrange solver in Section 6.8, incorporating also the variational density in the inner Lagrange solver. Moreover, the general theory of rotational invariant operators from Section 3.3 is discussed and the Kanamori parametrization of the Slater-Condon interaction is derived. We find that a non-zero multipole factor $1/\gamma$ pushes the phase boundary between the Mott insulator and the split-valence insulator down to lower J/U, enhancing the tendency to form the Hund's metal for even lower values of J/U. Our findings show that the full Coulomb interaction contain a number of aspects that is not captured by the Kanamori interaction.

Paper VI – Valence-skipping and negative-U in the d-band from repulsive local Coulomb interaction

Paper VI investigates a subtle aspect of the multipole interaction in the full Slater-Condon form of the Coulomb interaction. A study of the atomic ground state multiplet energy contributions reveal that for finite $1/\gamma$ a charge disproportionate phase separation occur for a range of J/U. In this range the d-band occupancy becomes valence-skipping, i.e. $d^n \rightarrow d^{n-1} + d^{n+1}$, but only for the particular multipole active d-band fillings d^1, d^4, d^6, and d^9. The boundaries for the phase separation is derived and the effect on the paramagnetic metal is investigated.

We find a novel valence-skipping metallic phase in direct contact with the Hund's metal, displaying qualitatively different charge fluctuations. The emergence of valence-skipping is driven by a negative effective U, known to drive charge order and anomalous superconductivity. The results are related to experimental observations of valence-skipping in iron d^4 and gold d^9 compounds. We also note that the multipole active fillings correlate with all current known families of anomalous superconductors, the d^9 cuprates, the d^4 ruthanates, and the d^6 iron pnictide and chalcogenides.

Bibliography

[1] H. Strand, *Critical Properties of the Mott-Hubbard Metal-Insulator Transition*, University of Gothenburg, 2011.

[2] P. Coleman, *The Evolving Monograph: Introduction to Many Body Physics*, 2009.

[3] F. Duan and J. Guojun, *Introduction to Condensed Matter Physics, vol 1.* (World Scientific Publishing Co. Pte. Ltd., 2005).

[4] J. Kohanoff, *Electronic Structure Calculations for Solids and Molecules* (Cambridge University Press, 2006).

[5] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[6] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[7] C. Filippi, C. J. Umrigar, and M. Taut, J. Chem. Phys. **100**, 1290 (1994).

[8] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).

[9] A. Georges, L. de' Medici, and J. Mravlje, Strong electronic correlations from Hund's coupling, arXiv:1207.3033, 2012.

[10] Armtuk, Periodic table based on html table from wikipedia, 2007.

[11] J. Hubbard, Proc. R. Soc. Lon. Ser.-A **276**, 238 (1963).

[12] R. M. Martin, *Electronic Structure Basic Theory and Practial Methods* (Cambridge University Press, 2004).

[13] M. Fabrizio, Sissa lecture notes, 2009.

[14] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Westview Press, 1998).

[15] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, Phys. Rev. B **52**, R5467 (1995).

[16] P. Limelette *et al.*, Science **302**, 89 (2003).

[17] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[18] D. J. Scalapino, Rev. Mod. Phys. **84**, 1383 (2012).

[19] K. Haule and G. Kotliar, New J. Phys. **11**, 025021 (2009).

[20] Y.-Q. Liang, N.-l. Di, and Z.-h. Cheng, Phys. Rev. B **72**, 134416 (2005).

[21] K. Kuzushita, S. Morimoto, S. Nasu, and S. Nakamura, J. Phys. Soc. Japan **69**, 2767 (2000).

[22] P. Fazekas, *Lecture Notes on Electron Correlation and Magnetism* (World Scientific Publishing Co. Pte. Ltd., 1999).

[23] A. Atland and B. Simons, *Condensed Matter Field Theory* (Cambridge University Press, 2006).

[24] C. Castellani, C. D. Castro, D. Feinberg, and J. Ranninger, Phys. Rev. Lett. **43**, 1957 (1979).

[25] N.-H. Tong, S.-Q. Shen, and F.-C. Pu, Phys. Rev. B **64**, 235109 (2001).

[26] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Dover Publications Inc., 2003).

[27] G. D. Mahan, *Many-Particle Physics*, 2nd ed. (Kluwer Academic/Plenum Publishers, 2000).

[28] N. F. Mott, Proc. Phys. Soc. A **62**, 416 (1949).

[29] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).

[30] Y. H. Szczech, M. A. Tusch, and D. E. Logan, Phys. Rev. Lett. **74**, 2804 (1995).

[31] R. Staudt, M. Dzierzawa, and A. Muramatsu, Eur. Phys. J. B **17**, 411 (2000).

[32] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, 1958).

[33] M. Golubitsky, SIAM Rev. **20**, 352 (1978).

[34] F. Lechermann *et al.*, Phys. Rev. B **74**, 125120 (2006).

[35] B. Amadon *et al.*, Phys. Rev. B **77**, 205112 (2008).

[36] J. Schirmer and L. S. Cederbaum, Phys. Rev. A **16**, 1575 (1977).

[37] L. Vaugier, H. Jiang, and S. Biermann, Phys. Rev. B **86**, 165105 (2012).

[38] B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill Book Company, 1963).

[39] I. Lindgren and J. Morrison, *Atomic Many-Body Theory*, 2nd ed. (Springer-Verlag, 1986).

[40] Z. Rudzikas, *Theoretical Atomic Spectroscopy*, Cambridge monographs on atomic, molecular and chemical physics (Cambridge University Press, 2007).

[41] J. F. Cornwell, *Group Theory in Physics: an Introduction* (Academic Press Inc., 1997).

[42] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (The Johns Hopkins Press, Baltimore, 1967).

[43] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, 1959).

[44] G. Racah, Phys. Rev. **63**, 367 (1943).

[45] M. Haverkort, *Spin and orbital degrees of freedom in transition metal oxides and oxide thin films studied by soft x-ray absorption spectroscopy*, PhD thesis, der Universität zu Köln, 2005.

[46] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963).

[47] T. Mizokawa and A. Fujimori, Phys. Rev. B **51**, 12880 (1995).

[48] N. Parragh, A. Toschi, K. Held, and G. Sangiovanni, Phys. Rev. B **86**, 155158 (2012).

[49] A. M. Läuchli and P. Werner, Phys. Rev. B **80**, 235117 (2009).

[50] H. Ishida and A. Liebsch, Phys. Rev. B **81**, 054513 (2010).

[51] A. Liebsch, Phys. Rev. B **84**, 180505 (2011).

[52] J. Kuneš *et al.*, Phys. Rev. Lett. **109**, 117206 (2012).

[53] N. Lanatà *et al.*, Phys. Rev. B **87**, 045122 (2013).

[54] G. Racah, Phys. Rev. **62**, 438 (1942).

[55] Y. Tanabe and S. Sugano, J. Phys. Soc. Jpn. **9**, 766 (1954).

[56] F. M. F. de Groot, J. C. Fuggle, B. T. Thole, and G. A. Sawatzky, Phys. Rev. B **42**, 5459 (1990).

[57] B. R. Judd and G. M. S. Lister, J. Phys. B: At. Mol. Phys. **17**, 3637 (1984).

[58] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyżyk, and G. A. Sawatzky, Phys. Rev. B **48** (1993).

[59] D. van der Marel and G. A. Sawatzky, Phys. Rev. B **37**, 10674 (1988).

[60] G. Kotliar *et al.*, Rev. Mod. Phys. **78**, 865 (2006).

[61] E. Müller-Hartmann, Z. Phys. B Con. Mat. **74**, 507 (1989).

[62] E. Müller-Hartmann, Z. Phys. B Con. Mat. **76**, 211 (1989).

[63] M. T. Heath, *Scientific Computing An Introductory Survey*, 2nd ed. (McGraw-Hill Book Company, 2002).

[64] R. Zitko, Phys. Rev. B **80**, 125125 (2009).

[65] G. Kotliar, E. Lange, and M. J. Rozenberg, Phys. Rev. Lett. **84**, 5180 (2000).

[66] C. G. Broyden, Math. Comput. **19**, 577 (1965).

[67] N. Blümer, *Mott-Hubbard Metal-Insulator Transition and Optical Conductivity in High Dimension*, PhD thesis, Universität Augsburg, 2002.

[68] E. Gull, *Continous-Time Quantum Monte Carlo Algorithms for Fermions*, PhD thesis, ETH Zürich, 2008.

[69] E. Gull *et al.*, Rev. Mod. Phys. **83**, 349 (2011).

[70] M. J. Rozenberg, X. Y. Zhang, and G. Kotliar, Phys. Rev. Lett. **69**, 1236 (1992).

[71] A. Georges and W. Krauth, Phys. Rev. B **48**, 7167 (1993).

[72] M. J. Rozenberg, G. Kotliar, and X. Y. Zhang, Phys. Rev. B **49**, 10181 (1994).

[73] P. W. Anderson, Phys. Rev. **124**, 41 (1961).

[74] K. Yosida and K. Yamada, Prog. Theor. Phys. Supp. **46**, 244 (1970).

[75] K. Yamada, Prog. Theor. Phys. **53**, 970 (1975).

[76] K. Yosida and K. Yamada, Prog. Theor. Phys. **53**, 1286 (1975).

[77] K. Yamada, Prog. Theor. Phys. **54**, 316 (1975).

[78] H. Kajueter and G. Kotliar, Phys. Rev. Lett. **77**, 131 (1996).

[79] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[80] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994).

[81] E. Koch, G. Sangiovanni, and O. Gunnarsson, Phys. Rev. B **78**, 115102 (2008).

[82] C. J. Bolech, S. S. Kancharla, and G. Kotliar, Phys. Rev. B **67**, 075110 (2003).

[83] M. Capone, M. Civelli, S. S. Kancharla, C. Castellani, and G. Kotliar, Phys. Rev. B **69**, 195105 (2004).

[84] D. Sénéchal, Phys. Rev. B **81**, 235125 (2010).

[85] H. Fehske, R. Schneider, and A. Weiße, *Computational Many-Particle Physics*, Lecture Notes in Physics Vol. 739 (Springer-Verlag Berlin Heidelberg, 2008).

[86] C. A. Perroni, H. Ishida, and A. Liebsch, Phys. Rev. B **75**, 045125 (2007).

[87] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, 2nd ed. (Manchester University Press, 1992).

[88] M. Capone, L. de'Medici, and A. Georges, Phys. Rev. B **76**, 245116 (2007).

[89] A. Liebsch, H. Ishida, and J. Merino, Phys. Rev. B **78**, 165123 (2008).

[90] G. van Rossum, CWI Report No. CS-R9526, 1995 (unpublished).

[91] E. Jones *et al.*, Scipy: Open source scientific tools for Python, 2001–.

[92] P. Peterson, Int. J. Comput. Sci. Eng. **4**, 296 (2009).

[93] R. Bulla, T. A. Costi, and T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008).

[94] P. Grete, S. Schmitt, C. Raas, F. B. Anders, and G. S. Uhrig, Phys. Rev. B **84**, 205104 (2011).

[95] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).

[96] M. C. Gutzwiller, Phys. Rev. **134**, A923 (1964).

[97] M. C. Gutzwiller, Phys. Rev. **137**, A1726 (1965).

[98] J. Bünemann, W. Weber, and F. Gebhard, Phys. Rev. B **57**, 6896 (1998).

[99] J. Bünemann, F. Gebhard, and W. Weber, Found. Phys. **30**, 2011 (2000).

[100] J. Bünemann *et al.*, Europhys. Lett. **61** (2003).

[101] J. Bünemann, F. Gebhard, T. Ohm, S. Weiser, and W. Weber, Frontiers in Magnetic Materials , 117 (2005).

[102] T. Schickling, F. Gebhard, and J. Bünemann, J.unemann, Phys. Rev. Lett. **106**, 146402 (2011).

[103] X. Deng, X. Dai, and Z. Fang, Europhys. Lett. **83** (2008).

[104] X. Deng, L. Wang, X. Dai, and Z. Fang, Phys. Rev. B **79**, 075114 (2009).

[105] J. Bünemann, F. Gebhard, T. Schickling, and W. Weber, Phys. Status Solidi B **249**, 1282 (2012).

[106] F. Lechermann, A. Georges, G. Kotliar, and O. Parcollet, Phys. Rev. B **76**, 155102 (2007).

[107] G. Kotliar and A. E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).

[108] J. Bünemann and F. Gebhard, Phys. Rev. B **76**, 193104 (2007).

[109] M. Fabrizio, Phys. Rev. B **76**, 165110 (2007).

[110] N. Lanatà, P. Barone, and M. Fabrizio, Phys. Rev. B **78**, 155127 (2008).

[111] N. Lanatà, P. Barone, and M. Fabrizio, Phys. Rev. B **80**, 224524 (2009).

[112] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, 1994).

[113] G. C. Wick, Phys. Rev. **80**, 268 (1950).

[114] H. Yokoyama and H. Shiba, J. Phys. Soc. Japan **56**, 1490 (1987).

[115] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, Phys. Rev. Lett. **94**, 026406 (2005).

[116] I. Peschel, Braz. J. Phys. **42**, 267 (2012).

[117] N. Lanatà, *The Gutzwiller variational approach to correlated systems*, PhD thesis, International School for Advanced Studies (SISSA/ISAS), 2009.

[118] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).

[119] W.-S. Wang *et al.*, Phys. Rev. B **82**, 125105 (2010).

[120] D. Kraft, ACM Transactions on Mathematical Software **20**, 262 (1994).

[121] M. Lax, *Symmetry Principles in Solid State and Molecular Physics* (Dover Publications Inc., 2001).

[122] G. F. Koster, Phys. Rev. **109**, 227 (1958).

[123] A. Togo, Spglib, http://spglib.sourceforge.net.

[124] The Atomic Simulation Environment (ASE), https://wiki.fysik.dtu.dk/ase.

[125] Bilbao crystallographic server, condensed matter physics dept. of the university of the basque country, http://www.cryst.ehu.es/.

[126] W. Metzner, Z. Phys. B **77**, 253 (1989).

Acknowledgments

This thesis, although written by the author, is not the product of a single person. As Newton pointed out, we all stand on the shoulders of giants, especially in physics. Therefore I would like to mention some people who have, sometimes unknowingly, carried me on their shoulders during my years as a Ph.D. student.

First of all I would like to thank my supervisor Prof. Bo Hellsing for believing in me; your continuous support and encouragement has been invaluable. My two closest collaborators have also been very important, my assistant supervisor Prof. Mats Granath; thank you for your patience with all my confused questions, and Nicola Lanatà; your energy and drive is truly inspiring. I also enjoyed working and discussing with Prof. Stellan Östlund and Andro Sabashvili in my first project. My examiner Prof. Henrik Johannesson also deserves mentioning, for always reassuring me that this thesis would turn out alright. I am grateful to Matteo Bazzanella and Anders Jonsson for careful reading of parts of the thesis.

My room mates, Matteo Bazzanella, Erik Eriksson, and Anders Ström have filled the days at work with inspiring discussions and a good laugh now and then. Also all current and former inhabitants of the third floor in Soliden has contributed to a inspiring and fun work environment. Dan Fors and Aleksandra Vojvodic deserve special mention for taking me under their wings during my first years. Apart from research I had a fun time working with the people at C3SE, fiddling with computer clusters and teaching Python to innocent fellow Ph.D. students, special thanks goes to Mattias Slabanja for patient mentoring.

My ten years in Gothenburg (mostly spent in the Physics building) have been enriched by many good friends Markus Billeter, Magnus Röding, Anders Jonsson, Viktor Griph, Per Magnusson, Andreas Skyman, Sara Sahlin, and Christina Wallgren, some who still linger at the university pursuing their own doctoral degrees. Finally I would like to thank my family for their continuous support, especially my father and siblings.

Appendix A

d^N multiplets

The d-electron Slater-Condon multiplets for all fillings are listed in Tables A.1, A.2, A.3, A.4, and A.5, in terms of the scaled Slater parameters F_k.

Table A.1: d^5 multiplets.

d^5					
^2S	$10F_0$	$-$	$3F_2$	$-$	$195F_4$
^2P	$10F_0$	$+$	$20F_2$	$-$	$240F_4$
^2D	$10F_0$	$-$	$4F_2$	$-$	$120F_4$
	$10F_0$	$-$	$3F_2$	$-$	$90F_4$
					$\pm 3\sqrt{57F_2^2 - 500F_2F_4 + 2300F_4^2}$
^2F	$10F_0$	$-$	$17F_2$	$-$	$90F_4$
					$\pm\sqrt{(8F_2 - 75F_4)^2}$
^2G	$10F_0$	$-$	$5F_2$	$-$	$150F_4$
					$\pm\sqrt{(8F_2 - 5F_4)^2}$
^2H	$10F_0$	$-$	$22F_2$	$-$	$30F_4$
^2I	$10F_0$	$-$	$24F_2$	$-$	$90F_4$
^4P	$10F_0$	$-$	$28F_2$	$-$	$105F_4$
^4D	$10F_0$	$-$	$18F_2$	$-$	$225F_4$
^4F	$10F_0$	$-$	$13F_2$	$-$	$180F_4$
^4G	$10F_0$	$-$	$25F_2$	$-$	$190F_4$
^6S	$10F_0$	$-$	$35F_2$	$-$	$315F_4$

Table A.2: d^0, d^1, d^2, d^3 multiplets.

d^0						
^1S	0					
d^1						
^2D	0					
d^2						
^1S	F_0	$+$	$14F_2$	$+$	$126F_4$
^1D	F_0	$-$	$3F_2$	$+$	$36F_4$
^1G	F_0	$+$	$4F_2$	$+$	F_4
^3P	F_0	$+$	$7F_2$	$-$	$84F_4$
^3F	F_0	$-$	$8F_2$	$-$	$9F_4$
d^3						
^2P	$3F_0$	$-$	$6F_2$	$-$	$12F_4$
^2D	$3F_0$	$+$	$5F_2$	$+$	$3F_4$
			$\pm\sqrt{193F_2^2 - 1650F_2F_4 + 8325F_4^2}$		
^2F	$3F_0$	$+$	$9F_2$	$-$	$87F_4$
^2G	$3F_0$	$-$	$11F_2$	$+$	$13F_4$
^2H	$3F_0$	$-$	$6F_2$	$-$	$12F_4$
^4P	$3F_0$	$-$	$147F_4$		
^4F	$3F_0$	$-$	$15F_2$	$-$	$72F_4$

Table A.3: d^{10}, d^9, d^8, d^7 multiplets.

d^{10}						
^1S	$45F_0$	$-$	$70F_2$	$-$	$630F_4$
d^9						
^2D	$36F_0$	$-$	$56F_2$	$-$	$504F_4$
d^8						
^1S	$28F_0$	$-$	$28F_2$	$-$	$252F_4$
^1D	$28F_0$	$-$	$45F_2$	$-$	$342F_4$
^1G	$28F_0$	$-$	$38F_2$	$-$	$377F_4$
^3P	$28F_0$	$-$	$35F_2$	$-$	$462F_4$
^3F	$28F_0$	$-$	$50F_2$	$-$	$387F_4$
d^7						
^2P	$21F_0$	$-$	$34F_2$	$-$	$264F_4$
^2D	$21F_0$	$-$	$23F_2$	$-$	$249F_4$
			$\pm\sqrt{193F_2^2 - 1650F_2F_4 + 8325F_4^2}$		
^2F	$21F_0$	$-$	$19F_2$	$-$	$339F_4$
^2G	$21F_0$	$-$	$39F_2$	$-$	$239F_4$
^2H	$21F_0$	$-$	$34F_2$	$-$	$264F_4$
^4P	$21F_0$	$-$	$28F_2$	$-$	$399F_4$
^4F	$21F_0$	$-$	$43F_2$	$-$	$324F_4$

Table A.4: d^4 multiplets.

d^4	
${}^1S$	$6F_0 + 10F_2 + 6F_4 \pm 2\sqrt{193F_2^2 - 1650F_2F_4 + 8325F_4^2}$
${}^1D$	$6F_0 + 9F_2 - \frac{153}{2}F_4 \pm \frac{3}{2}\sqrt{144F_2^2 - 1160F_2F_4 + 3425F_4^2}$
${}^1F$	$6F_0 - 84F_4$
${}^1G$	$6F_0 - 5F_2 - \frac{13}{2}F_4 \pm \frac{1}{2}\sqrt{708F_2^2 - 7500F_2F_4 + 30825F_4^2}$
${}^1I$	$6F_0 - 15F_2 - 9F_4$
${}^3P$	$6F_0 - 5F_2 - \frac{153}{2}F_4 \pm \frac{1}{2}\sqrt{912F_2^2 - 9960F_2F_4 + 38025F_4^2}$
${}^3D$	$6F_0 - 5F_2 - 129F_4$
${}^3F$	$6F_0 - 5F_2 - \frac{153}{2}F_4 \pm \frac{3}{2}\sqrt{68F_2^2 - 540F_2F_4 + 2225F_4^2}$
${}^3G$	$6F_0 - 12F_2 - 94F_4$
${}^3H$	$6F_0 - 17F_2 - 69F_4$
${}^5D$	$6F_0 - 21F_2 - 189F_4$

Table A.5: d^6 multiplets.

d^6	
${}^1S$	$15F_0 - 4F_2 - 120F_4 \pm 2\sqrt{193F_2^2 - 1650F_2F_4 + 8325F_4^2}$
${}^1D$	$15F_0 - 5F_2 - \frac{405}{2}F_4 \pm \frac{3}{2}\sqrt{144F_2^2 - 1160F_2F_4 + 3425F_4^2}$
${}^1F$	$15F_0 - 14F_2 - 210F_4$
${}^1G$	$15F_0 - 19F_2 - \frac{265}{2}F_4 \pm \frac{1}{2}\sqrt{708F_2^2 - 7500F_2F_4 + 30825F_4^2}$
${}^1I$	$15F_0 - 29F_2 - 135F_4$
${}^3P$	$15F_0 - 19F_2 - \frac{405}{2}F_4 \pm \frac{1}{2}\sqrt{912F_2^2 - 9960F_2F_4 + 38025F_4^2}$
${}^3D$	$15F_0 - 19F_2 - 255F_4$
${}^3F$	$15F_0 - 19F_2 - \frac{405}{2}F_4 \pm \frac{3}{2}\sqrt{68F_2^2 - 540F_2F_4 + 2225F_4^2}$
${}^3G$	$15F_0 - 26F_2 - 220F_4$
${}^3H$	$15F_0 - 31F_2 - 195F_4$
${}^5D$	$15F_0 - 35F_2 - 315F_4$

Appendix B

Gutzwiller Approximation

In this appendix we re-derive the results of the Gutzwiller approximation, which was only stated as fact in Chapter 6. The goal is to make it clear why the evaluation of expectation values with respect to the Gutzwiller wave-function Ψ_G is in itself a many-body problem, and to show how Eqs. (6.10) and (6.11) arise from the constraints GA1 and GA2 [Eqs. (6.8) and (6.9)] and the limit of infinite dimensions. Finally, the relations for the renormalization factors and effective hopping of Eqs. (6.13) and (6.14) are derived.

B.1 Wick's theorem

The main weapon in the calculation of expectation values is the fact that Ψ_0 is a Slater determinant wave-function. For this class of wave-functions we can use Wick's theorem [113]. The theorem states that the expectation value of a product of fermion operators with respect to Ψ_0 is given by the sum of all combinations of paired quadratic expectation values.

First we introduce some notation, by denoting expectation values with respect to Ψ_0 as $\langle\Psi_0|\cdot|\Psi_0\rangle = \langle\cdot\rangle_0$, and defining the fermion operators $q_{\mathbf{R}a}$ as to span all local creation and annihilation operators $c_{\mathbf{R}\alpha}$ and $c^\dagger_{\mathbf{R}\beta}$

$$q_{\mathbf{R}a} = \{c_{\mathbf{R}1}, \ldots, c_{\mathbf{R}m}; c^\dagger_{\mathbf{R}1}, \ldots, c^\dagger_{\mathbf{R}m}\}\,, \tag{B.1}$$

where m is the number of single-particle fermion states. Thereby $\langle q^\dagger_{\mathbf{R}a} q_{\mathbf{R}b}\rangle_0$ includes all possible combinations

$$\langle c^\dagger_{\mathbf{R}\alpha} c_{\mathbf{R}\beta}\rangle_0\,, \langle c_{\mathbf{R}\alpha} c^\dagger_{\mathbf{R}\beta}\rangle_0\,, \langle c_{\mathbf{R}\alpha} c_{\mathbf{R}\beta}\rangle_0\,, \text{ and } \langle c^\dagger_{\mathbf{R}\alpha} c^\dagger_{\mathbf{R}\beta}\rangle_0\,.$$

As an example of Wick's theorem, consider a quartic term of fermion operators. All possible contractions (i.e. pairings), denoted by lines connecting operators, can then be expressed as

$$\langle q_1 q_2 q_3 q_4 \rangle_0 = \langle \overbrace{q_1}\, \overbrace{q_2}\, \overbrace{q_3}\, \overbrace{q_4} \rangle_0 + \langle \overbrace{q_1 \overbrace{q_2}\, \overbrace{q_3}\, q_4} \rangle_0 + \langle \overbrace{q_1 \overbrace{q_2}\, \overbrace{q_3}\, \overbrace{q_4}} \rangle_0 = \\ \langle q_1 q_2 \rangle_0 \langle q_3 q_4 \rangle_0 + \langle q_1 q_4 \rangle_0 \langle q_2 q_3 \rangle_0 + \langle q_1 q_3 \rangle_0 \langle q_2 q_4 \rangle_0 \quad \text{(B.2)}$$

Generalizing to a term of order $2n$, it can be written as the sum over all ordered pairs

$$\langle q_{a_1} \dots q_{a_{2n}} \rangle_0 = \sum_P \langle q_{b_{P_1}} q_{c_{P_1}} \rangle_0 \dots \langle q_{b_{P_n}} q_{c_{P_n}} \rangle_0 \,, \quad \text{(B.3)}$$

where the sum goes over all pairings P conserving order, i.e. for every set of pair indices $b_{P_l} = a_k$, $c_{P_l} = a_m$, $k < m$.

It will also turn out to be useful to write the expectation value of a general operator product in terms of Wick contractions. Such a product takes the form of a sum of all possible orders p of inter-operator contractions

$$\langle \hat{A}\hat{B} \rangle_0 = \langle \hat{A} \rangle_0 \langle \hat{B} \rangle_0 + \sum_{p=1}^{\infty} \langle \overbrace{\hat{A} \dots \hat{B}}^{\times p} \rangle_0 = \langle \hat{A} \rangle_0 \langle \hat{B} \rangle_0 + \langle \overbrace{\hat{A}\hat{B}} \rangle_0 + \langle \overbrace{\overbrace{\hat{A}\hat{B}}} \rangle_0 + \dots \,, \quad \text{(B.4)}$$

between the operators $\hat{A}$ and $\hat{B}$, that in turn can be any combination of fermion $q_{\mathbf{R}\alpha}$ operators.

If the two operators act exclusively on two different sites $\mathbf{R}$ and $\mathbf{R}'$, the contractions can formally be written as

$$\langle \overbrace{\hat{A}_{\mathbf{R}} \hat{B}_{\mathbf{R}'}} \rangle_0 = \sum_{ab} \langle \hat{A}^{\bar{a}}_{\mathbf{R}} \rangle_0 \langle \hat{B}^{\bar{b}}_{\mathbf{R}'} \rangle_0 \langle q_{\mathbf{R}a} q_{\mathbf{R}'b} \rangle_0 \quad \text{(B.5)}$$

$$\langle \overbrace{\overbrace{\hat{A}_{\mathbf{R}} \hat{B}_{\mathbf{R}'}}} \rangle_0 = \sum_{abcd} \langle \hat{A}^{\bar{a}\bar{b}}_{\mathbf{R}} \rangle_0 \langle \hat{B}^{\bar{c}\bar{d}}_{\mathbf{R}'} \rangle_0 \langle \overbrace{\overbrace{\hat{Q}_{\mathbf{R}ab}\ \hat{Q}}}_{\mathbf{R}'cd} \rangle_0 \,, \quad \text{(B.6)}$$

using the notation $\hat{A}^{\bar{a}}_{\mathbf{R}}$ for an operator $\hat{A}_{\mathbf{R}}$ with a fermion operator $q_{\mathbf{R}a}$ removed, and the generalized single-particle density operator $\hat{Q}_{\mathbf{R}ab} = q_{\mathbf{R}a} q_{\mathbf{R}b}$. By now we are well prepared to appreciate the problem of evaluating expectation values of the Gutzwiller wave-function Ψ_G.

B.2 Gutzwiller expectation values

Let us at this point recall the definition of the Gutzwiller wave-function Ψ_G in Eqs. (6.1) to (6.3), and write down the complete expectation value with respect to a local operator $\hat{\mathcal{O}}(\mathbf{R})$

$$\langle\Psi_G|\hat{\mathcal{O}}(\mathbf{R})|\Psi_G\rangle = \langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}\Big(\prod_{\mathbf{R}'\neq\mathbf{R}}\hat{\mathcal{P}}^\dagger_{\mathbf{R}'}\hat{\mathcal{P}}_{\mathbf{R}'}\Big)|\Psi_0\rangle\,. \tag{B.7}$$

That this is not directly equal to $\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}|\Psi_0\rangle$ as indicated in Eq. (6.10) is clear from Wick's theorem and Section B.1. An exact evaluation must also account for the contractions between the operators at $\mathbf{R}$ and all projectors acting at the other sites $\mathbf{R}'$.

To handle the explosion of possible contractions we rewrite the product of projector pairs $\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}}$ in terms of the associated operator $\hat{\mathcal{W}}_{\mathbf{R}} = \hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}} - 1$, which gives

$$\prod_{\mathbf{R}}\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{P}}_{\mathbf{R}} = \prod_{\mathbf{R}}\Big(1+\hat{\mathcal{W}}_{\mathbf{R}}\Big) = 1 + \sum_{m=1}^{\infty}\frac{1}{m!}{\sum_{\mathbf{R}_1,\ldots,\mathbf{R}_m}}' \hat{\mathcal{W}}_{\mathbf{R}_1}\cdots\hat{\mathcal{W}}_{\mathbf{R}_m}\,, \tag{B.8}$$

where $\mathbf{R}_1,\ldots,\mathbf{R}_m$ are all different in the primed sum. Thus, the local expectation value can be written as the sum

$$\begin{aligned}\langle\Psi_G|\hat{\mathcal{O}}(\mathbf{R})|\Psi_G\rangle &= \langle\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}\rangle_0 \\ &\quad + \sum_{m=1}^{\infty}\frac{1}{m!}{\sum_{\mathbf{R}_1\ldots\mathbf{R}_m\neq\mathbf{R}}}' \langle\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}\hat{\mathcal{W}}_{\mathbf{R}_1}\ldots\hat{\mathcal{W}}_{\mathbf{R}_m}\rangle_0\,.\end{aligned} \tag{B.9}$$

As $\hat{\mathcal{P}}^\dagger_{\mathbf{R}}\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{P}}_{\mathbf{R}}$ is a local operator we can, for the sake of brevity, and without loss of generality, denote it as a general local operator $\hat{\mathcal{O}}$. Now, for the anticipated result [Eq. (6.10)] to hold only the first term should survive in the Gutzwiller approximation. To prove this we start by studying the first order term in terms of its Wick contractions

$$\sum_{\mathbf{R}'\neq\mathbf{R}}\langle\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}\rangle_0 = \langle\hat{\mathcal{O}}(\mathbf{R})\rangle_0\sum_{\mathbf{R}'\neq\mathbf{R}}\langle\hat{\mathcal{W}}_{\mathbf{R}'}\rangle_0 + \sum_{\mathbf{R}'\neq\mathbf{R}}\sum_{p=1}^{\infty}\langle\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}^{\times p}\rangle_0\,. \tag{B.10}$$

and show how each order of contractions vanishes.

The zeroth order contraction in Eq. (B.10) is zero by virtue of GA1 [Eq. (6.8)] and the definition of $\hat{\mathcal{W}}_{\mathbf{R}}$.

$$1 = \langle \hat{\mathcal{P}}_{\mathbf{R}}^{\dagger} \hat{\mathcal{P}}_{\mathbf{R}} \rangle_0 = 1 + \langle \hat{\mathcal{W}}_{\mathbf{R}} \rangle_0 \,, \quad \Rightarrow \quad \langle \hat{\mathcal{W}}_{\mathbf{R}} \rangle_0 = 0 \,. \tag{B.11}$$

This also causes all disconnected contractions of $\hat{\mathcal{W}}_{\mathbf{R}}$ to disappear at higher orders in Eq. (B.9).

B.3 First and second order contractions

To understand why the first order contraction – and all odd order contractions for that matter – are zero, it is instructive to return to the definition of the projector $\hat{\mathcal{P}}_{\mathbf{R}}$ in Eq. (6.4) and construct the corresponding expression for $\hat{\mathcal{W}}_{\mathbf{R}}$

$$\hat{\mathcal{W}}_{\mathbf{R}} = \sum_{\Gamma\Gamma'} w_{\Gamma\Gamma'} \; |\Gamma, \mathbf{R}\rangle\langle\Gamma', \mathbf{R}| \tag{B.12}$$

where the coefficients $w_{\Gamma\Gamma'}$ are determined by the configuration weights $\lambda_{\Gamma\Gamma'}$ of $\hat{\mathcal{P}}_{\mathbf{R}}$ in Eq. (6.4)

$$w_{\Gamma\Gamma'} = -\delta_{\Gamma\Gamma'} + \sum_{\Gamma''} \lambda^*_{\Gamma''\Gamma} \lambda_{\Gamma''\Gamma'} \tag{B.13}$$

Before performing any Wick contractions we first express $\hat{\mathcal{W}}_{\mathbf{R}}$ in terms of fermion operators. This is achieved by using the definition of the local occupation number states in Eq. (6.5) and noting that the projector onto the site local vacuum state $|0, \mathbf{R}\rangle\langle 0, \mathbf{R}|$ can be written as

$$|0, \mathbf{R}\rangle\langle 0, \mathbf{R}| = \prod_{\alpha} (1 - \hat{n}_{\alpha\mathbf{R}}) \,. \tag{B.14}$$

Thus, $\hat{\mathcal{W}}_{\mathbf{R}}$ can be expressed as

$$\hat{\mathcal{W}}_{\mathbf{R}} = \sum_{\Gamma\Gamma'} w_{\Gamma\Gamma'}(\mathbf{R}) \prod_{\alpha\in\Gamma} c^{\dagger}_{\alpha\mathbf{R}} \prod_{\beta} (1 - \hat{n}_{\beta\mathbf{R}}) \prod_{\gamma\in\Gamma} c_{\gamma\mathbf{R}} \,. \tag{B.15}$$

We already know that the projector, and thus $\hat{\mathcal{W}}_{\mathbf{R}}$, conserve particle number, which in turn guarantees that there is an *even* total number of fermion operators in every term of Eq. (B.15). This causes the expectation value of $\hat{\mathcal{W}}_{\mathbf{R}}$ with an odd number of operators q_a removed to disappear

$$\langle \hat{\mathcal{W}}^{\bar{a}}_{\mathbf{R}} \rangle_0 = 0 \,, \quad \langle \hat{\mathcal{W}}^{\bar{a}\bar{b}\bar{c}}_{\mathbf{R}} \rangle_0 = 0 \,, \text{ etc..} \tag{B.16}$$

as in each case there is an unpaired fermion operator that gives a factor $\langle q_b \rangle_0 = 0$ in all Wick contractions. These terms occurs of course in every odd order p of Wick contractions, hence

$$\langle \hat{\mathcal{O}} \overbrace{\hat{\mathcal{W}}_{\mathbf{R}}}^{p} \rangle_0 = 0\,, \quad \forall p = 2n+1\,. \tag{B.17}$$

The zeroth and first order contractions can now be used to understand the consequences of GA2 that gives the relation

$$\begin{aligned}\langle q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 &= \langle \hat{\mathcal{P}}^\dagger_{\mathbf{R}} \hat{\mathcal{P}}_{\mathbf{R}} q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 = \langle q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 + \langle \hat{\mathcal{W}}_{\mathbf{R}} q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 \\ &= \langle q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 + \langle \hat{\mathcal{W}}_{\mathbf{R}} \rangle_0 \langle q_{\mathbf{R}a} q_{\mathbf{R}b} \rangle_0 + \langle \overbrace{\hat{\mathcal{W}}_{\mathbf{R}} \hat{Q}_{\mathbf{R}ab}} \rangle_0 + \langle \overbrace{\overbrace{\hat{\mathcal{W}}_{\mathbf{R}} \hat{Q}_{\mathbf{R}ab}}} \rangle_0\,, \end{aligned} \tag{B.18}$$

where on the last line the second term is zero due to GA1 and Eq. (B.11), and the third term is an odd contraction and zero due to Eq. (B.17). In other words the second order contraction of $\hat{\mathcal{W}}_{\mathbf{R}}$ with the local single-particle density operator $\hat{Q}_{\mathbf{R}ab} = q_{\mathbf{R}a} q_{\mathbf{R}b}$ is also zero

$$\langle \overbrace{\overbrace{\hat{\mathcal{W}}_{\mathbf{R}} \hat{Q}_{\mathbf{R}ab}}} \rangle_0 = 0\,, \tag{B.19}$$

for any a and b, and any Slater determinant Ψ_0, i.e. for any values of $\langle q_c q_d \rangle$. Performing the second order contraction explicitly

$$\begin{aligned}\langle \overbrace{\overbrace{\hat{\mathcal{W}}_{\mathbf{R}} \hat{Q}_{\mathbf{R}ab}}} \rangle_0 &= \sum_{cd} \langle \hat{\mathcal{W}}^{\bar{c}\bar{d}}_{\mathbf{R}} \rangle_0 \langle \overbrace{\hat{Q}_{\mathbf{R}cd} \hat{Q}_{\mathbf{R}ab}} \rangle_0 \\ &= \sum_{cd} \langle \hat{\mathcal{W}}^{\bar{c}\bar{d}}_{\mathbf{R}} \rangle_0 \left(\langle q_c q_a \rangle_0 \langle q_d q_b \rangle_0 + \langle q_c q_b \rangle_0 \langle q_d q_a \rangle_0 \right)\,, \end{aligned} \tag{B.20}$$

the only way this can be zero is that the expectation value of $\hat{\mathcal{W}}_{\mathbf{R}}$ with two fermion operators removed is identically equal to zero

$$\langle \hat{\mathcal{W}}^{\bar{a}\bar{b}}_{\mathbf{R}} \rangle_0 = 0\,. \tag{B.21}$$

This has consequences far beyond local contractions as this is exactly the factor that appear in every second order contraction of $\hat{\mathcal{W}}$

$$\langle \overbrace{\hat{\mathcal{O}}_{\mathbf{R}} \hat{\mathcal{W}}_{\mathbf{R}'}} \rangle_0 = \sum_{abcd} \langle \hat{\mathcal{O}}^{\bar{a}\bar{b}}_{\mathbf{R}} \rangle_0 \langle \hat{\mathcal{W}}^{\bar{c}\bar{d}}_{\mathbf{R}'} \rangle_0 \langle \overbrace{\hat{Q}_{\mathbf{R}ab} \hat{Q}_{\mathbf{R}'cd}} \rangle_0 = 0\,. \tag{B.22}$$

Thus GA2 actually causes all second order contractions of $\hat{\mathcal{W}}_{\mathbf{R}}$ to be zero.

So what we have left now, after treating the zeroth, the second, and all odd order contractions, are the even contractions of fourth order and higher. To handle them we first need to introduce the limit of infinite dimensions.

B.4 Infinite dimensional limit

The infinite dimensional limit ($d \to \infty$) is a powerful starting point for constructing non-trivial theories for interacting electrons. In our case, for the Gutzwiller wave-function Ψ_G, the limit simplifies Wick contractions considerably due to the dimensional scaling of the single-particle density matrix $\langle \Psi_0 | c^\dagger_{\mathbf{R}\alpha} c_{\mathbf{R}'\beta} | \Psi_0 \rangle$.

Following Ref. [126], consider for example a single-band Slater determinant which obeys the innocent looking relation[1]

$$\sum_{\mathbf{R}'} |\langle \Psi_0 | c^\dagger_{\mathbf{R}} c_{\mathbf{R}'} | \Psi_0 \rangle|^2 = \langle \Psi_0 | c^\dagger_{\mathbf{R}} c_{\mathbf{R}} | \Psi_0 \rangle = n^0 \,. \tag{B.23}$$

The sum over $\mathbf{R}'$ can be partitioned over all sites at New York distance[2] l (i.e. the l_1 norm $l = |\mathbf{R} - \mathbf{R}'|_1$) that belong to the same irreducible representation j of the point group. Using the short hand notation $|\langle \Psi_0 | c^\dagger_{\mathbf{R}} c_{\mathbf{R}'} | \Psi_0 \rangle| = \mathcal{C}_{\mathbf{R}'\mathbf{R}} = \mathcal{C}_j$ for these correlators, Eq. (B.23) takes the form

$$\sum_{l=1}^{\infty} \sum_{j} n_j(l) \mathcal{C}_j^2 = n^0 - (n^0)^2 \,, \tag{B.24}$$

where $n_j(l)$ is the number of sites belonging to the irreducible representation at distance l. This directly gives an upper bound on the correlators

$$|\langle \Psi_0 | c^\dagger_{\mathbf{R}} c_{\mathbf{R}'} | \Psi_0 \rangle| \leq \sqrt{\frac{n^0(1-n^0)}{n_j(l)}} \propto \frac{1}{d^{l/2}} \,, \tag{B.25}$$

as $n_j(l)$ scales like $n_j(l) \propto d^l$. The derivation can directly be generalized to multi-band Slater determinants using orbital averages, and in the end the same scaling relation is obtained for the general correlators $\langle \Psi_0 | c^\dagger_{\mathbf{R}\alpha} c_{\mathbf{R}'\beta} | \Psi_0 \rangle$. Applied to a sum of correlators to some integer power p the scaling gives

$$\sum_{\mathbf{R}' \neq \mathbf{R}} \left(\langle \Psi_0 | c^\dagger_{\mathbf{R}\alpha} c_{\mathbf{R}'\beta} | \Psi_0 \rangle \right)^p \propto \sum_{l=1}^{\infty} d^l (d^{-\frac{l}{2}})^p \propto \sum_{l=1}^{\infty} d^{(1-\frac{p}{2})l} \,. \tag{B.26}$$

Thus for powers $p > 2$, the sum goes to zero as $d \to \infty$. This property causes the sum of all p-order Wick contractions of an operator $\hat{\mathcal{O}}$ at site $\mathbf{R}$

[1] Obtained using the idempotence of $\langle \Psi_0 | c^\dagger_{\mathbf{k}} c_{\mathbf{k}} | \Psi_0 \rangle$, $\langle \Psi_0 | c^\dagger_{\mathbf{k}} c_{\mathbf{k}} | \Psi_0 \rangle = \langle \Psi_0 | c^\dagger_{\mathbf{k}} c_{\mathbf{k}} | \Psi_0 \rangle^2$.

[2] The know-it-alls call this the Manhattan distance. The author is going the long distance to New York to find it, but only after finishing this thesis.

with another operator $\hat{\mathcal{W}}_{\mathbf{R}'}$ on every other site to vanish for $p > 2$

$$\sum_{\mathbf{R}'\neq\mathbf{R}} \langle\Psi_0|\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}^{\times p}|\Psi_0\rangle \xrightarrow[d\to\infty]{} 0\,,\ \text{iff } p > 2\,. \tag{B.27}$$

Returning to the first order term in Eq. (B.10) of the expansion in Eq. (B.9), we can now treat all orders of contractions

$$\sum_{\mathbf{R}'\neq\mathbf{R}} \langle\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}\rangle_0 = \sum_{\mathbf{R}'\neq\mathbf{R}} \Big[\langle\hat{\mathcal{O}}(\mathbf{R})\rangle_0\langle\hat{\mathcal{W}}_{\mathbf{R}'}\rangle_0 + \langle\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}\rangle_0$$

$$+ \langle\overbrace{\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}}\rangle_0 + \langle\overbrace{\overbrace{\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}}}\rangle_0 + \sum_{p=4}^{\infty}\langle\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}'}}^{\times p}\rangle_0\Big] \xrightarrow[d\to\infty]{} 0\,, \tag{B.28}$$

where the zeroth order contraction term is zero due to GA1, the first and third order due to number conservation, the second order due to GA2, and finally all orders ≥ 4 are zero due to the limit of infinite dimensions.

In the same way we can treat all higher order terms, where there is no need to keep track of the explicit contraction structure, only the contraction order p_n of each $\hat{\mathcal{W}}_{\mathbf{R}_n}$ term matters

$$\sum_{\mathbf{R}_1\ldots\mathbf{R}_m}{}' \sum_{p_0\ldots p_m} \langle\overbrace{\hat{\mathcal{O}}(\mathbf{R})\hat{\mathcal{W}}_{\mathbf{R}_1}\ldots\hat{\mathcal{W}}_{\mathbf{R}_m}}^{p_0\ \ p_1\ \ \ p_m}\rangle_0\,. \tag{B.29}$$

If any of the contraction orders $p_1, \ldots p_n$ are less than 3, the contribution is identically equal to zero, due to GA1, GA2 and number conservation. When all orders $p_1, \ldots p_n$ are greater than 4, the expectation value scales as $d^{-p/2}$ where p is the total number of contractions given by

$$p = \frac{1}{2}\sum_m p_m \geq \frac{p_0 + 4m}{2} \tag{B.30}$$

and the sum scales as d^m giving the over all scaling $d^{-p_0/2}$. Thus the connected contraction with $p_0 = 2$ goes to zero, but the disconnected contraction does not! However, all is not lost. Just like in perturbation theory prevailing disconnected contractions (i.e. diagrams) are a reminder that we need to properly normalize our wave-function. This is accomplished by computing the expectation values on the form

$$\frac{\langle\Psi_G|\hat{\mathcal{O}}(\mathbf{R})|\Psi_G\rangle}{\langle\Psi_G|\Psi_G\rangle}\,, \tag{B.31}$$

where by the linked cluster theorem [14], all disconnected terms in the nominator are canceled by the denominator. Defining $\{\dots\}_0^c$ as all fully connected contractions, we have thus established that to all orders the connected terms vanish

$$\sum_{m=1}^{\infty} \frac{1}{m!} \sum_{\mathbf{R}_1 \dots \mathbf{R}_m}{}' \{\hat{\mathcal{O}}(\mathbf{R}) \hat{\mathcal{W}}_{\mathbf{R}_1} \dots \hat{\mathcal{W}}_{\mathbf{R}_m}\}_0^c \xrightarrow[d\to\infty]{} 0 \tag{B.32}$$

Returning to the correctly normalized expectation value of a local operator with respect to Ψ_G, it is exactly given by

$$\begin{aligned} \frac{\langle \Psi_G | \hat{\mathcal{O}}(\mathbf{R}) | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} &= \langle \hat{\mathcal{P}}_{\mathbf{R}}^\dagger \hat{\mathcal{O}}(\mathbf{R}) \hat{\mathcal{P}}_{\mathbf{R}} \rangle_0 \\ &\quad + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{\mathbf{R}_1 \dots \mathbf{R}_m}{}' \{\hat{\mathcal{P}}_{\mathbf{R}}^\dagger \hat{\mathcal{O}}(\mathbf{R}) \hat{\mathcal{P}}_{\mathbf{R}} \hat{\mathcal{W}}_{\mathbf{R}_1} \dots \hat{\mathcal{W}}_{\mathbf{R}_m}\}_0^c \end{aligned} \tag{B.33}$$

So, by imposing GA1, GA2 and taking the limit of infinite dimensions the expectation value simplifies to

$$\frac{\langle \Psi_G | \hat{\mathcal{O}}(\mathbf{R}) | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} = \langle \hat{\mathcal{P}}_{\mathbf{R}}^\dagger \hat{\mathcal{O}}(\mathbf{R}) \hat{\mathcal{P}}_{\mathbf{R}} \rangle_0 \,. \tag{B.34}$$

Along the same lines the quadratic non-local expectation value reduces to

$$\frac{\langle \Psi_G | c_{\mathbf{R}\alpha}^\dagger c_{\mathbf{R}'\beta} | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} = \langle [\hat{\mathcal{P}}_{\mathbf{R}}^\dagger c_{\mathbf{R}\alpha}^\dagger \hat{\mathcal{P}}_{\mathbf{R}}] [\hat{\mathcal{P}}_{\mathbf{R}'}^\dagger c_{\mathbf{R}'\beta} \hat{\mathcal{P}}_{\mathbf{R}'}] \rangle_0 \,, \tag{B.35}$$

under the same assumptions.

B.5 Renormalized hopping

The quadratic hopping expression in Eq. (B.35) can be simplified further as indicated in Eq. (6.12). To derive this relation we write the non-local quadratic expectation value of Eq. (B.35) as

$$\langle \Psi_0 | [\hat{\mathcal{P}}_{\mathbf{R}}^\dagger c_{\mathbf{R}\alpha}^\dagger \hat{\mathcal{P}}_{\mathbf{R}}] [\hat{\mathcal{P}}_{\mathbf{R}'}^\dagger c_{\mathbf{R}'\beta} \hat{\mathcal{P}}_{\mathbf{R}'}] | \Psi_0 \rangle = \langle \Psi_0 | \hat{\mathcal{R}}_{\mathbf{R}\alpha}^\dagger \hat{\mathcal{R}}_{\mathbf{R}'\beta} | \Psi_0 \rangle \tag{B.36}$$

by introducing the compound operator $\hat{\mathcal{R}}_{\mathbf{R}\alpha} = \hat{\mathcal{P}}_{\mathbf{R}}^\dagger c_{\mathbf{R}\alpha} \hat{\mathcal{P}}_{\mathbf{R}}$. Now the Wick decomposition of Eq. (B.36) contains all orders of contractions between these

operators

$$\langle\Psi_0|\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{R}}_{\mathbf{R}'\beta}|\Psi_0\rangle = \langle\Psi_0|\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}|\Psi_0\rangle\langle\Psi_0|\hat{\mathcal{R}}_{\mathbf{R}'\beta}|\Psi_0\rangle + \langle\Psi_0|\overbrace{\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{R}}_{\mathbf{R}'\beta}}|\Psi_0\rangle + \langle\Psi_0|\overbrace{\overbrace{\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{R}}_{\mathbf{R}'\beta}}}|\Psi_0\rangle + \langle\Psi_0|\overbrace{\overbrace{\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{R}}_{\mathbf{R}'\beta}}}|\Psi_0\rangle + \dots\,, \quad \text{(B.37)}$$

with each line corresponding to a Wick contraction joining a pair of fermion operators on sites $\mathbf{R}$ and $\mathbf{R}'$, respectively. Assuming that the projector $\hat{\mathcal{P}}_{\mathbf{R}}$ is number conserving, only containing terms with an even number of fermion operators, then $\hat{\mathcal{R}}_{\mathbf{R}\alpha}$ is guaranteed to only have odd terms. This causes all even number of contractions to be zero due to unpaired fermion operators. Terms with three or more fermionic lines joining $\mathbf{R}$ and $\mathbf{R}'$ on the other hand disappears in the limit of infinite dimensions. The only remaining term is the single contraction, which can be explicitly written in terms of the operators $\hat{\mathcal{R}}^{\bar{b}}_{\mathbf{R}'\beta}$, obtained from $\hat{\mathcal{R}}_{\mathbf{R}'\beta}$ after removal of a $q_{\mathbf{R}'b}$ fermion operator

$$\langle\Psi_0|\overbrace{\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{R}}_{\mathbf{R}'\beta}}|\Psi_0\rangle = \sum_{ab}\langle\Psi_0|\hat{\mathcal{R}}^{\bar{a}\dagger}_{\mathbf{R}\alpha}|\Psi_0\rangle\langle\Psi_0|\hat{\mathcal{R}}^{\bar{b}}_{\mathbf{R}'\beta}|\Psi_0\rangle\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}'b}|\Psi_0\rangle \quad \text{(B.38)}$$

Thus, quadratic expectation values of the form $\langle c^\dagger_{\mathbf{R}}c_{\mathbf{R}'}\rangle_G$ are related to the corresponding Slater determinant expectation values $\langle q^\dagger_{\mathbf{R}}q_{\mathbf{R}'}\rangle_0$ through a renormalization provided by $\langle\hat{\mathcal{R}}^{\bar{b}}_{\mathbf{R}'\beta}\rangle_0$. Fortunately, $\langle\hat{\mathcal{R}}^{\bar{b}}_{\mathbf{R}'\beta}\rangle_0$ occurs also in other Wick contractions, such as

$$\begin{aligned}&\langle\Psi_0|[\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{P}}_{\mathbf{R}}]q_{\mathbf{R}b}|\Psi_0\rangle = \langle\Psi_0|\hat{\mathcal{R}}^\dagger_{\mathbf{R}\alpha}q_{\mathbf{R}b}|\Psi_0\rangle \\ &= \sum_a\langle\Psi_0|\hat{\mathcal{R}}^{\bar{a}\dagger}_{\mathbf{R}\alpha}|\Psi_0\rangle\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}b}|\Psi_0\rangle = \sum_a\mathcal{R}_{a\alpha}(\mathbf{R})\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}b}|\Psi_0\rangle \quad \text{(B.39)}\end{aligned}$$

where we have introduced the shorthand notation $\mathcal{R}_{a\alpha}(\mathbf{R}) = \langle\Psi_0|\hat{\mathcal{R}}^{\bar{a}\dagger}_{\mathbf{R}\alpha}|\Psi_0\rangle$ for the renormalization factors. In other words, non-local quadratic expectation values can be written as

$$\langle\Psi_G|c^\dagger_{\mathbf{R}\alpha}c_{\mathbf{R}'\beta}|\Psi_G\rangle = \sum_{ab}\mathcal{R}_{a\alpha}(\mathbf{R})\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}'b}|\Psi_0\rangle\mathcal{R}^\dagger_{\beta b}(\mathbf{R}')\,. \quad \text{(B.40)}$$

Using this we can rewrite the kinetic energy $\langle\hat{\mathcal{T}}\rangle_G$ of the Gutzwiller wavefunction in Eq. (6.6) directly in terms of Slater determinant expectation

values and the renormalization factors $\mathcal{R}$ as

$$\begin{aligned}\langle\Psi_G|\hat{\mathcal{T}}|\Psi_G\rangle &= \sum_{\mathbf{R}\mathbf{R}'}\sum_{\alpha\beta} t^{\alpha\beta}_{\mathbf{R}\mathbf{R}'}\langle\Psi_G|c^\dagger_{\mathbf{R}\alpha}c_{\mathbf{R}'\beta}|\Psi_G\rangle \\ &= \sum_{\mathbf{R}\mathbf{R}'}\sum_{ab}\left[\sum_{\alpha\beta}\mathcal{R}_{a\alpha}(\mathbf{R})t^{\alpha\beta}_{\mathbf{R}\mathbf{R}'}\mathcal{R}^\dagger_{\beta b}(\mathbf{R}')\right]\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}'b}|\Psi_0\rangle \\ &= \sum_{\mathbf{R}\mathbf{R}'}\sum_{ab}\tilde{t}^{ab}_{\mathbf{R}\mathbf{R}'}\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}'b}|\Psi_0\rangle = \langle\Psi_0|\hat{\mathcal{T}}^G[\mathcal{R}]|\Psi_0\rangle\,, \quad \text{(B.41)}\end{aligned}$$

where $\hat{\mathcal{T}}^G$ is an effective kinetic Hamiltonian

$$\hat{\mathcal{T}}^G[\mathcal{R}] = \sum_{\mathbf{R}\mathbf{R}'}\sum_{ab}\tilde{t}^{ab}_{\mathbf{R}\mathbf{R}'}q^\dagger_{\mathbf{R}a}q_{\mathbf{R}'b}\,, \quad \text{(B.42)}$$

with hopping amplitudes

$$\tilde{t}^{ab}_{\mathbf{R}\mathbf{R}'} = \sum_{\alpha\beta}\mathcal{R}_{a\alpha}(\mathbf{R})t^{\alpha\beta}_{\mathbf{R}\mathbf{R}'}\mathcal{R}^\dagger_{\beta b}(\mathbf{R}')\,, \quad \text{(B.43)}$$

renormalized by $\mathcal{R}$. The renormalization factors $\mathcal{R}$ can be calculated from purely local expectation values through

$$\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{P}}_{\mathbf{R}}q_{\mathbf{R}b}|\Psi_0\rangle = \sum_a\mathcal{R}_{a\alpha}(\mathbf{R})\langle\Psi_0|q^\dagger_{\mathbf{R}a}q_{\mathbf{R}b}|\Psi_0\rangle\,. \quad \text{(B.44)}$$

Note that $q^\dagger_{\mathbf{R}a}$ span both $c^\dagger_{\mathbf{R}\alpha}$ and $c_{\mathbf{R}\alpha}$, so inserting the creation and annihilation operators explicitly still gives the original expression of Fabrizio [109]

$$\begin{aligned}&\langle\Psi_0|\hat{\mathcal{P}}^\dagger_{\mathbf{R}}c^\dagger_{\mathbf{R}\alpha}\hat{\mathcal{P}}_{\mathbf{R}}c_{\mathbf{R}\beta}|\Psi_0\rangle \\ &\quad= \sum_\gamma\mathcal{R}_{\gamma\alpha}(\mathbf{R})\langle\Psi_0|c^\dagger_{\mathbf{R}\gamma}c_{\mathbf{R}\beta}|\Psi_0\rangle + \sum_\gamma\Delta_{\gamma\alpha}(\mathbf{R})\langle\Psi_0|c_{\mathbf{R}\gamma}c_{\mathbf{R}\beta}|\Psi_0\rangle \quad \text{(B.45)}\end{aligned}$$

where $\Delta_{\gamma\alpha}$ is the last part of $\mathcal{R}_{a\alpha}$, $\Delta_{\gamma\alpha} = \mathcal{R}_{(n+\gamma)\alpha}$, for n local fermions, as $1 \leq \alpha \leq n$ and $1 \leq a \leq 2n$. Further, $\Delta_{\gamma\alpha}$ is only non-zero when Ψ_0 is a BCS type of wave-function with non-vanishing fermion pair creation and annihilation expectation values $\langle c^\dagger c^\dagger\rangle_0 \neq 0$. Note that this case is not treated in Chapter 6, but is a straightforward generalization.

Appendix C

On the method of Lagrange multipliers

In this appendix we outline the essential steps required to use the method of Lagrange multipliers for constrained optimization problems in complex variables with real valued target functions and complex valued equality constraints. The approach where the complex variable z and its conjugate $\bar{z}$ are considered as independent variables is explained and the method of complex Lagrange multipliers for complex valued constraints is re-derived.

C.1 Lagrange formulation

For a general real valued minimization problem with an equality constraint

$$\min_x \{f(x) : g(x) = c\}\,, \tag{C.1}$$

the method of Lagrange multipliers provides a strategy to find feasible stationary points. The corresponding Lagrange function is defined as

$$\mathcal{L}(x, \lambda) = f(x) + \lambda(g(x) - c)\,. \tag{C.2}$$

A necessary condition for optimality of a solution is that $\nabla\mathcal{L} = \mathbf{0}$. Thus, solving the root problem in x and λ

$$\nabla\mathcal{L}[x,\lambda] = \mathbf{0} \quad \Rightarrow \quad \begin{cases} \dfrac{\partial\mathcal{L}}{\partial x} = f'(x) + \lambda g'(x) = 0 \\ \dfrac{\partial\mathcal{L}}{\partial\lambda} = g(x) - c = 0 \end{cases}, \tag{C.3}$$

gives feasible stationary solutions of the minimization problem.

C.2 Real functions and complex variables

A real-valued function $\mathcal{L}$ in one complex variable $z \in \mathbb{C}$, $\mathcal{L} : \mathbb{C} \to \mathbb{R}$ can equivalently be written as a function of two independent real variables $\mathcal{L} : \mathbb{R}^2 \to \mathbb{R}$, by letting $z = x + iy$. The condition for a zero gradient is then

$$\nabla\mathcal{L}(x,y) = \mathbf{0} \quad \Rightarrow \quad \begin{cases} \dfrac{\partial\mathcal{L}}{\partial x} = 0 \\ \dfrac{\partial\mathcal{L}}{\partial y} = 0 \end{cases}. \tag{C.4}$$

Making a change of variables from x and y to the complex linearly independent z and $\bar{z}$

$$\begin{cases} z = x + iy \\ \bar{z} = x - iy \end{cases} \quad \Rightarrow \quad \begin{cases} x = \dfrac{1}{2}(z + \bar{z}) \\ y = \dfrac{1}{2i}(z - \bar{z}) \end{cases}, \tag{C.5}$$

gives another set of partial derivatives, with respect to either z or $\bar{z}$. Using the chain rule the corresponding gradients takes the form

$$\left.\begin{aligned} \left.\frac{\partial\mathcal{L}}{\partial z}\right|_{\bar{z}} &= \frac{\partial x}{\partial z}\frac{\partial\mathcal{L}}{\partial x} + \frac{\partial y}{\partial z}\frac{\partial\mathcal{L}}{\partial y} = \frac{1}{2}\frac{\partial\mathcal{L}}{\partial x} - \frac{i}{2}\frac{\partial\mathcal{L}}{\partial y} \\ \left.\frac{\partial\mathcal{L}}{\partial\bar{z}}\right|_{z} &= \frac{\partial x}{\partial\bar{z}}\frac{\partial\mathcal{L}}{\partial x} + \frac{\partial y}{\partial\bar{z}}\frac{\partial\mathcal{L}}{\partial y} = \frac{1}{2}\frac{\partial\mathcal{L}}{\partial x} + \frac{i}{2}\frac{\partial\mathcal{L}}{\partial y} \end{aligned}\right\} \Rightarrow \frac{\partial\mathcal{L}}{\partial\bar{z}} = \left(\frac{\partial\mathcal{L}}{\partial z}\right)^*. \tag{C.6}$$

Thus as $\mathcal{L}$ is real-valued it suffices to impose either $\partial_z\mathcal{L} = 0$ or $\partial_{\bar{z}}\mathcal{L} = 0$ to fulfill the zero gradient condition.

C.3 Complex variables and constraints

Consider the optimization problem

$$\min_z \{f(z) : g(z) = c\}\,, \tag{C.7}$$

where $z, c \in \mathbb{C}$, $f : \mathbb{C} \to \mathbb{R}$ and $g : \mathbb{C} \to \mathbb{C}$. From the previous section we know that the complex variable poses no problem, and we turn our attention to the complex constraint. By separating the constraint in real and imaginary parts we get the problem

$$\min_z \{f(z) : g_R(z) = c_R, g_I(z) = c_I\}\,, \tag{C.8}$$

using the notation $g_R(z) = \mathrm{Re}[g(z)]$, $g_I(z) = \mathrm{Im}[g(z)]$. The two real valued constraints can of course be imposed separately by using two Lagrange multipliers by the Lagrange function

$$\mathcal{L}(z, \lambda_R, \lambda_I) = f(z) + \lambda_R(g_R(z) - c_R) + \lambda_I(g_I(z) - c_I)\,, \tag{C.9}$$

and using the two previous sections the necessary gradient equations for an optimal solution are

$$\left.\frac{\partial \mathcal{L}}{\partial \bar{z}}\right|_z = 0\,, \quad \frac{\partial \mathcal{L}}{\partial \lambda_R} = g_R(x, y) - c_R = 0\,, \quad \frac{\partial \mathcal{L}}{\partial \lambda_I} = g_I(x, y) - c_I = 0\,. \tag{C.10}$$

Instead of considering the real and imaginary parts of the constraint separately it is possible to use a complex Lagrange multiplier λ and rewrite the constraints using λ and $\bar{\lambda}$. By defining $\lambda \equiv \lambda_R - i\lambda_I$ it is straightforward to show that

$$\lambda_R(g_R(z) - c_R) + \lambda_I(g_I(z) - c_I) = \frac{1}{2}\lambda(g(z) - z) + \frac{1}{2}\bar{\lambda}(\bar{g}(z) - \bar{c})\,. \tag{C.11}$$

Thus the equivalent Lagrangian function with one complex Lagrange multiplier is given by

$$\mathcal{L}(z, \lambda) = f(z) + \frac{1}{2}\lambda(g(z) - c) + \frac{1}{2}\bar{\lambda}(\bar{g}(z) - \bar{c})\,. \tag{C.12}$$

The factor $1/2$ can of course be absorbed in the Lagrange multiplier.

Finally the gradient equations for our Lagrange function in one complex variable and one complex Lagrange multiplier now takes the deceptively simple form

$$\left.\frac{\partial \mathcal{L}}{\partial \bar{z}}\right|_z = 0\,, \quad \left.\frac{\partial \mathcal{L}}{\partial \lambda}\right|_{\bar{\lambda}} = 0\,. \tag{C.13}$$

In what follows we will keep the factor $1/2$ in the Lagrange function in order to simplify the treatment of the case with a real constraint ($g(z), c \in \mathbb{R}$). In that case also $\lambda \in \mathbb{R}$ and we regain the original form of the Lagrange multiplier theorem

$$\mathcal{L}(z, \lambda) = f(z) + \lambda(g(z) - c)\,. \tag{C.14}$$

Paper I

The Dynamical Mean Field Theory phase space extension and critical properties of the finite temperature Mott transition

Hugo U. R. Strand, Andro Sabashvili, Mats Granath, Bo Hellsing, and Stellan Östlund

Phys. Rev. B **83**, 205136 (2011)

Paper II

Distributional exact diagonalization formalism for quantum impurity models

Mats Granath and Hugo U. R. Strand

Phys. Rev. B **86**, 115111 (2012)

Paper III

Efficient implementation of the Gutzwiller variational method

Nicola Lanatà, Hugo U. R. Strand, Xi Dai, and Bo Hellsing

Phys. Rev. B **85**, 035133 (2012)

Paper IV

Orbital Selectivity in Hund's metals: The Iron Chalcogenides

Nicola Lanatà, Hugo U. R. Strand, Gianluca Giovannetti, Bo Hellsing, Luca de' Medici, and Massimo Capone

Phys. Rev. B **87**, 045122 (2013)

Paper V

Local correlation in the d-band, Slater-Condon vs. Kanamori

Hugo U. R. Strand, Nicola Lanatà, Mats Granath, and Bo Hellsing

(in manuscript)

Paper VI

Valence skipping and effective negative-U in the d-band from repulsive local Coulomb interaction

Hugo U. R. Strand

(in manuscript)